UPPSALA UNIVERSITET

UUIC-B22-2

THE MOLECULAR STRUCTURE OF HUMAN CARBONIC ANHYDRASE, FORM C AND INHIBITOR COMPLEXES

by

IVAR VAARA

UPPSALA 1974

THE MOLECULAR STRUCTURE OF HUMAN CARBONIC ANHYDRASE, FORM C AND INHIBITOR COMPLEXES

BY

IVAR VAARA

CONTENTS

INTRODUCTION 1

METHODS 3

- Purification 3
- Crystallization and heavy atom derivatives 3
- Collection of X-ray data 4
- Heavy atom parameters and phase angles 10

RESULTS AND DISCUSSION 10

- The shape of the molecule 13
- Identification of the main chain 13
- Identification of side chains 30
- "Model building" 34

SECONDARY AND TERTIARY STRUCTURE 48

- Helical segments 48
- Turns 51
- β-structures 52
- Aspects of folding 55
- Location of various side chains 56
- Comparison with spectral studies 59

ACTIVE SITE 60

- The metal ion 60
- Various side chains in the active site 63
- The "water structure" 67

ANION BINDING TO HCAC 69

SULFONAMIDE BINDING TO HCAC 73

STRUCTURE STUDIES AT LOW pH 77

SUMMARY 78

ACKNOWLEDGEMENTS 80

REFERENCES 82

UNIV. LEIDEN BIBL.

INTRODUCTION

Early in this century the importance of enzymes in living systems was established. Later, when the structural homogeneity of proteins was recognized, it became possible to study enzymes in terms of the chemical groups which were involved in coenzyme and substrate binding and to begin to understand the molecular mechanisms of catalysis.

Among the many methods that can give useful information about protein structures are chemical sequence determination, kinetic studies, spectral studies and X-ray diffraction techniques. The X-ray diffraction method has proved to be uniquely powerful for the determination of detailed structures of proteins and of other macromolecules. So far, the structures of more than 25 proteins and details of the active sites of many enzymes have been determined at high resolution (2-3Å).

The high resolution structures of proteins have also elucidated the following phenomena:

a) "Induced fit" : conformational changes in the protein tertiary structures caused by the substrate binding may result in a more exact, functional "lock-and-key" fit between enzyme and substrate (5).

b) "Allosteric effects" : conformational changes in the quarternary structures of multisubunit proteins may lead to "cooperativity" between subunits (6).

c) "Gene multiplication" : analysis of the tertiary structure and amino acid sequence of different parts of a protein molecule has indicated that a polypeptide chain may be coded by a gene which has arisen by the multiplication of an ancestral gene (7).

d) "Antigen-antibody interaction" : elucidation of the structures

of antigen-antibody complexes is expected to provide a detailed understanding of biologically important antigen-antibody interactions (8).

This thesis describes the molecular structure of human carbonic anhydrase form C (HCAC) at 2.0 Å resolution and the complexes between the enzyme and various inhibitors including anions and sulfonamides. The comparison of HCAC and its isoenzyme HCAB is also discussed.

Carbonic anhydrase (EC 4. 2. 1. 1. $CO_2 + H_2O \rightleftarrows H^+ + HCO_3^-$) in human red blood cells was discovered in 1932 (10, 11). Later Keilin and Mann (12) found that the enzyme contained zinc. Carbonic anhydrase is found in many types of cells in plants, animals and bacteria (13, 14). In animals and bacteria it is a monomer. However, in parsley (14), in pea leaves (15) and in spinach (15) it is a hexamer, containing six zinc atoms per molecule. The spinach carbonic anhydrase is inhibited by anions and sulfonamides (16).

The enzyme consists in most mammals of about 259 amino acid residues with a molecular weight of about 30,000 (17) and contains one firmly bound zinc ion per molecule (12), which is essential for activity (18). It is inhibited by anions (10) and sulfonamides non-competitively (12). The sulfonamides constitute an important class of drugs, which are used as diuretics and for treatment of glaucoma (19). The chemical sequences of bovine and sheep carbonic anhydrases show large similarities with the human isoenzymes (20). Although there is over 60% sequence homology between the major forms of human carbonic anhydrase, HCAB and HCAC, they differ significantly in activity. HCAC is about 10 times faster than HCAB at neutral pH. However, HCAB approaches the activity of HCAC at higher pH (21). The turnover rate of HCAC reaches 10^6 sec^{-1} for carbon dioxide hydration at pH 8.5. Thus carbonic anhydrase is one of the fastest enzymes known. In vitro, carbonic anhydrase also catalyzes the hydration of aldehydes and the hydrolysis of esters (22).

The crystallization of HCAC was reported in 1962 (23), and the preparation of useful heavy atom derivatives in 1964 (24, 25). A low resolution structure was presented in 1967 (26). The high resolution structure was interpreted with a partial amino acid sequence in 1972 (1) and with the complete sequence in this thesis.

METHODS

Purification

Several different purification procedures (27,28, 1) have been used to try to obtain crystals suitable for the structure investigation. Most of the data have been recorded on crystals obtained using the method described by Liljas et al. (1). Carbonic anhydrase activity was measured according to Schneider and Liefländer (29).

Crystallization and heavy atom derivatives

Crystals obtained directly from native enzyme solution were flakey and unsuitable for high resolution (2. 0 Å) structure determination. However, crystals grown from solution of the enzyme reacted with 5 equivalents of methylmercurithioglycolic acid (MMTGA) (provided initially by the late Dr B. Tilander and subsequently by Prof. A. Fredga) were very suitable for high resolution work and good native crystals could be obtained by removing the mercurial by soaking in 0. 01 M cysteine (1). The HCAC-AMSulf (3-acetoxymercuri-4-aminobenzenesulfonamide) crystals were also obtained by co-crystallization. The dimension of the crystals were about 0. 20 x 0. 20 x 0. 35 mm. Four derivatives were used in the structure investigation of HCAC to high resolution. Two of them were complexes of HCAC with the above mentioned MMTGA and AMSulf, the remaining two were complexes

with iodide and goldcyanide obtained by soaking native enzyme crystals in 0.1M NH_4I and 0.2M $KAu(CN)_2$ solutions, respectively, for at least six days.

Collection of X-ray data

Human carbonic anhydrase form C crystallizes in space group $P2_1$ with the cell dimensions a=42.7, b=41.7, c=73.0 Å and $\beta=104.6^0$.

The X-ray data have been collected photographically using Buerger Supper precession cameras with Ni-filtered CuK_α-radiation from Philips fine-focus tubes. Thirtytwo sets of precession films were collected for the native enzyme and each derivative except for the iodide and goldcyanide derivatives which had significantly lower heavy atom contributions at higher resolution (1). For the latter two derivatives twentytwo layers covering almost all the reflexions with spacing greater than 2.5 Å were collected. Generally one crystal was used per layer. The films have been measured on a single beam, semi-automatic microdensitometer constructed by Dr. V. Klimecki, Institute of Chemistry, University of Uppsala (30).

The data of the bromide and acetazolamide (Diamox) complexes of HCAC were measured on a fully automatic microdensitometer also constructed by Dr. V. Klimecki in collaboration with our group. A short description of the apparatus and its function follows:

Two step-motors drive the planar film table in two perpendicular directions in steps of 10μ and 40μ. The light spot used is 40μ x 40μ. The distance between neighbouring reflexions is preset in such a way that every reflexion is within a rectangular area (fig. 1). A reference system defines the start and end of each row of reflexions. Another "built-in" system limits the integration to only that part of the film where reflexions are expected. The film is scanned in a zig-zag pattern (fig. 1) and the analogue integration is performed in one of two ways:

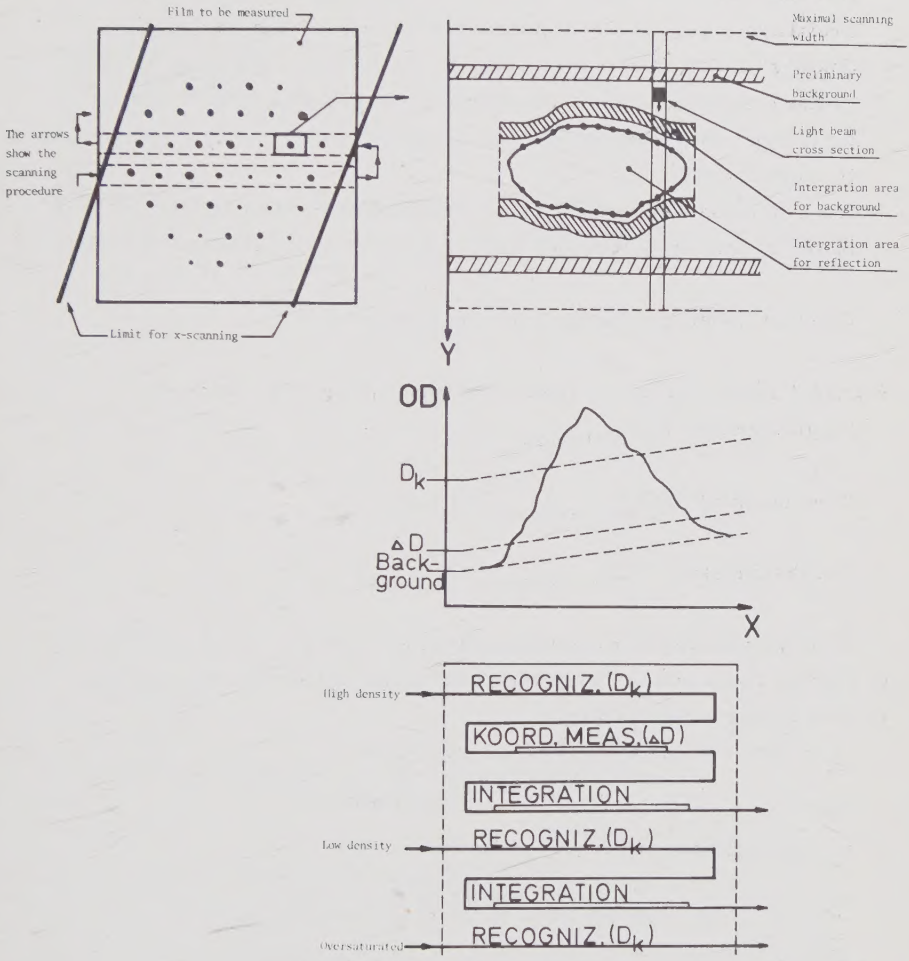


Fig. 1 The main principle of function of the automatic microdensitometer.

If a reflexion has a peak optical density higher than a preset arbitrary level, D_k, the area of the reflexion with an optical density higher than $\triangle D$ is defined. $\triangle D$ defines the start and end of the reflexion in the Y-direction for every 40 μ step in the X-direction. The integration area of the reflexion and the background is shown in fig. 1.
If a reflexion has a peak optical density lower than D_k the "profile" of the reflexion is taken as that of the last well determined reflexion.

Oversaturated reflexions are not integrated.

Figure 1 shows the above mentioned three programs named

"high density" ($OD_{reflexion} > D_k$),

"low density" ($OD_{reflexion} < D_k$) and

"oversaturated" ($OD_{reflexion} > OD_{max}$).

Every reflexion is measured within a "box" and an output record is written after every "box" which simplifies indexing. Each output record contains values for,

I_R = total optical density of the reflexion

I_B = " " " " " background

t_R = integration time over reflexion

t_B = " " " background

The observed intensity may then be calculated by :

$$I_{obs} = I_R - \frac{t_R}{t_B} \cdot I_B$$

A linearity correction is built into the instrument so that the apparatus is linear up to an optical density of about 2.2 (fig. 2). Figures 3 and 4 give information on the accuracy of the instrument.

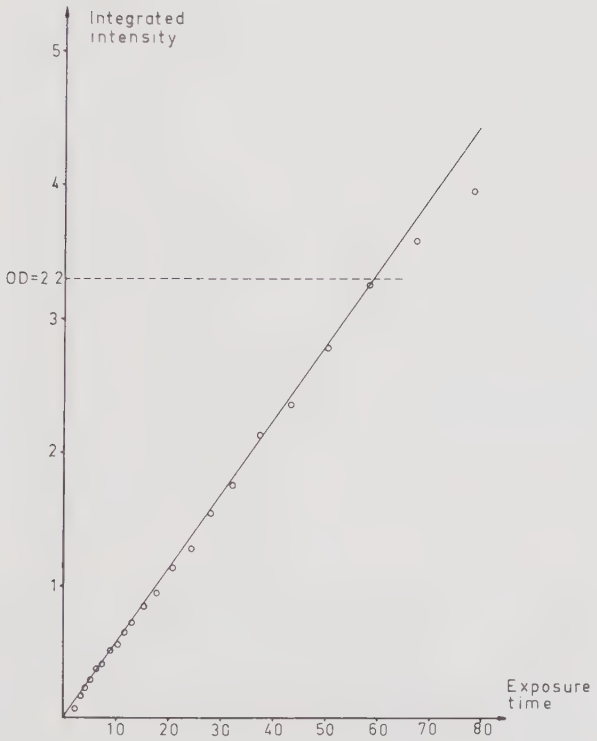


Fig. 2 Integrated intensity as a function of exposure time for a calibration film.

In the refinement and phase angle calculations discussed later the minimum I_{obs} used was 0.05 which is equal to an optical density of about 0.07.

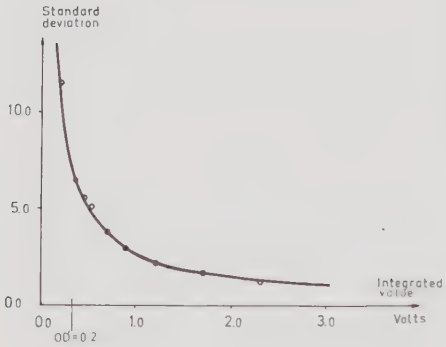


Fig. 3

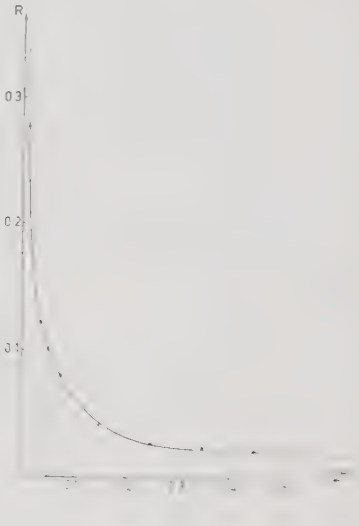


Fig. 4

Fig. 3 Standard deviation between integrated intensity values of pairs of symmetry-related reflexions as a function of integrated intensity. OD' = 0.2 is marked.
On the actual film number of reflexion pairs with

OD' < 0.2 was 340
OD' > 0.2 was 246
OD' is defined as $^{10}\log\frac{I_B}{I_R}$

Fig. 4 Scaling residual, R, for common reflexions of the native enzyme films in different intervals of I_{obs}, calculated by the HRS-program.

$$R = \frac{\text{scaling } \sum \underline{hkl} \cdot \sum_{i=1}^{n} |F_i - F_{av}|}{\text{scaling } \sum \underline{hkl} \cdot n \times F_{av}}$$

F_i = scaled amplitude for a given reflexion on one layer
F_{av} = average amplitude for the n measurements of reflexions with the same ($\underline{hkl}$) value.

Table 1. Heavy atom parameters refined within the 2.0 Å sphere with least squares methods.

DERIVATIVE	POSITION	Z	x	y	z	B	T	b_{11}	b_{22}	b_{33}	b_{12}	b_{23}	b_{13}	E	R_1	R_2
MMTGA	M_1	60.3	0.095	0.261	0.345		-0.07	2.1	2.3	1.8	1.3	2.4	2.9	27	0.08	0.55
AMSulf	A_1	56.0	0.103	0.257	0.355		0.07	1.9	1.0	2.6	0.5	1.4	4.3	25	0.07	0.33
	A_2	17.8	0.177	0.256	0.444			3.8	5.5	-0.2	0.3	-4.7	-1.3			
	A_3	32.8	-0.041	0.027	0.235			4.9	1.5	3.3	3.5	-4.0	1.2			
	A_4	61.8	-0.001	0.050	0.170			1.4	-0.2	0.3	0.4	1.0	-0.3			
	A_5	21.1	0.415	-0.101	0.330			2.4	2.0	0.5	-2.1	-2.1	2.0			
I^-	I_3	23.5	-0.041	0.027	0.244	22.5	1.0							32	0.10	0.82
	I_6	16.7	0.264	-0.186	0.314	13.7										
	I_7	4.2	0.026	0.176	0.193	9.9										
$Au(CN)_2^-$	G_8	65.4	-0.046	0.098	0.224		1.4	2.5	3.7	2.2	2.7	-4.7	1.3	32	0.10	0.61
Zn^a			-0.064	0.036	0.218											

$$R_1 = \frac{\Sigma \left| |F_H| - |F_p| \cdot e^{i\alpha} + f_c \right|}{\Sigma |F_H|}$$

$$R_2 = \frac{\Sigma \left| |F_H| - |F_p| \cdot e^{i\alpha} + f_c \right|}{\Sigma \left| |F_H| - |F_p| \right|}$$

a = obtained from the electron density maps
Z = occupancy of heavy atom site, electrons
B = isotropic temperature factors
T = overall temp. factors
b_{11}, b_{12} - - - = anisotropic temperature factors x 10^5
E = r.m.s error from lack of closure refinement
F_p, F_H and f_c are structure factors for the native enzyme, the heavy atom derivate and the calculated structure factors for the heavy atom contribution, respectively, α is the phase angle for the native enzyme.

Heavy atom parameters and phase angles

Preliminary least squares refinements were performed with projection data. The final refinement used full three dimensional data including about 11.500 reflexions. The refinement and phase angle calculations (31) were computed with a program kindly provided by Dr H. Muirhead (32) and modified by Dr K. K. Kannan.

In order to locate secondary sites difference Fourier maps were calculated. In the refinement anisotropic temperature factors were used for the MMTGA, AMSulf and $Au(CN)_2^-$ derivatives and an isotropic temperature factor was used for iodide. The refined heavy atom parameters are shown in table 1.

A "best" Fourier (33) synthesis was calculated in planes perpendicular to the x-axis. The maps were calculated and plotted using Dr G. N. Reeke's Fourier program which prints the map at the required scale. The electron density was traced onto Mylar sheets mounted on freely hanging aluminium frames.

The model was constructed with Kendrew-Watson skeletal model parts on the scale 2 cm = 1 Å, using a Richards optical device with a half silvered mirror (34), which enables the model and the electron density map to be viewed simultaneously. The arrangement is shown in fig. 5. The methods concerning HCAC have been described in more detail elsewhere (1, 2).

RESULTS AND DISCUSSION

The high resolution structure of HCAC, first interpreted on the basis of partial amino acid sequences, was published in the following publications:

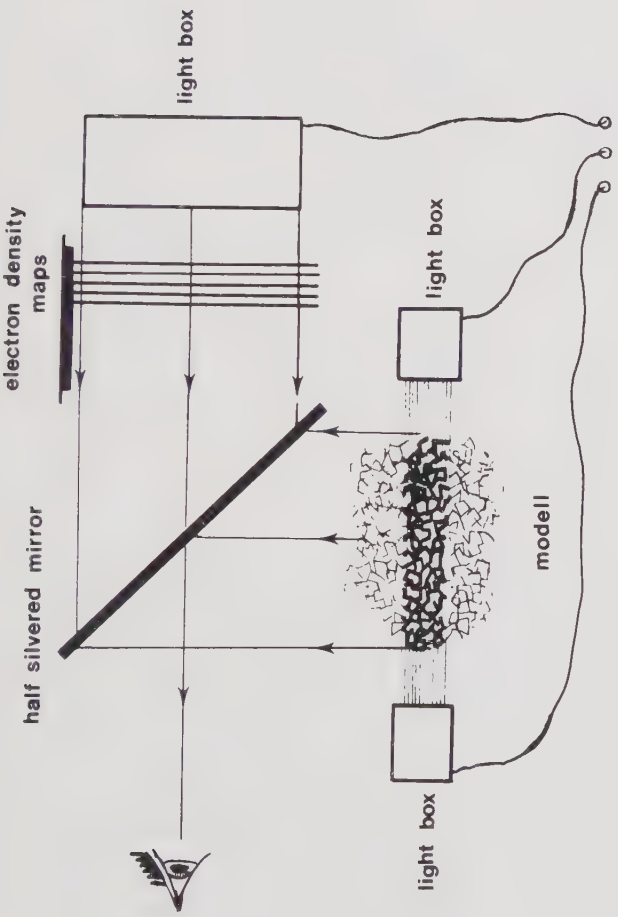


Fig. 5 A schematic drawing of the model building device.

Crystal Structure of Human Erythrocyte Carbonic Anhydrase C. VI. The Three-dimensional Structure at High Resolution in Relation to other Mammalian Carbonic Anhydrases. K. K. Kannan, A. Liljas, I. Vaara, P. C. Bergstén, S. Lövgren, B. Strandberg, U. Bengtsson, U. Carlbom, K. Fridborg, L. Järup, and M. Petef. Cold Spring Harbor Symposia on Quantitative Biology 36 (1971),221. (1)

Crystal Structure of Human Carbonic Anhydrase C. A. Liljas, K. K. Kannan, P.-C. Bergstén, I. Vaara, K. Fridborg, B. Strandberg, U. Carlbom, L. Järup, S. Lövgren, and M. Petef. Nature New Biology 235 (1972),131. (2)

Crystal Structure of Human Erythrocyte Carbonic Anhydrase C. V. Complexes with some Anion and Sulfonamide Inhibitors. P.-C. Bergstén, I. Vaara, S. Lövgren, A. Liljas, K. K. Kannan, and U. Bengtsson. Alfred Benzon Symposium IV (Rörth, M. and Astrup, D. eds.) Munksgaard p. 363 (1971). (3)

Functional Aspects of the Three-dimensional Structure of the Active Site of Carbonic Anhydrase. I. Vaara, S. Lövgren, A. Liljas, K. K. Kannan, and P.-C. Bergstén. The Second International Conference on Red Cell Metabolism and Function (Brewer, J. ed.) Ann Arbor, Vol. 28 (1972) p. 169. (4)

The assumption was made that the sequence of HCAC was homologous with that of the isoenzyme HCAB for which most of the sequence was known at that time (35). In many places where the sequence was unknown some residues could be identified from the shape of the electron density. For example, the peptide containing the zinc ligands had not been sequenced, but these ligands were interpreted to be histidyl residues, which was confirmed when the chemical sequence was established later.

The interpretation of the electron density described in the following is based on the complete chemical sequences of HCAC as determined by Henderson et al. (36).

Certain regions of the polypeptide chain were also revised with the help of the chemically established sequence of HCAC and a more satisfactory fit between the electron density map and the chemical sequence was obtained. Minor changes occured on the surface of the molecule between residues 100-105, 200-210, 233-245. The volume of the active site cavity increased due to the removal of the loop consisting of residues 200-210 and the topography of the active site is somewhat different with the new HCAC primary sequence.

The overall structural homology between HCAB and HCAC has also improved. Fig.a gives a comparative view of the HCAC and the HCAB molecules.

The shape of the molecule

The molecular boundary is well determined from the electron density map (2), and the molecule is about 40x42x55 Å measured between extreme points on the backbone. The molecule can be illustrated as a sailboat (fig. 19), where the bottom is defined by the residues 41-55, 150-159, and 174-187, the front is formed by the residues 165-170 and 229-238, and the back by residues 252-256 and 37-41, with rudder defined by the three residues in the C terminus. The deck is formed by the large pleated sheet, on the middle of which the zinc ion is situated in an open door cabin, formed by the residues 125-138, 194-204 and 243-251. Finally the sail is the 25 amino acid residues at the amino terminus.

Identification of the main chain

In general the main chain could be followed unambigously. More than 90 % of all carbonyl groups of the peptide bonds could easily

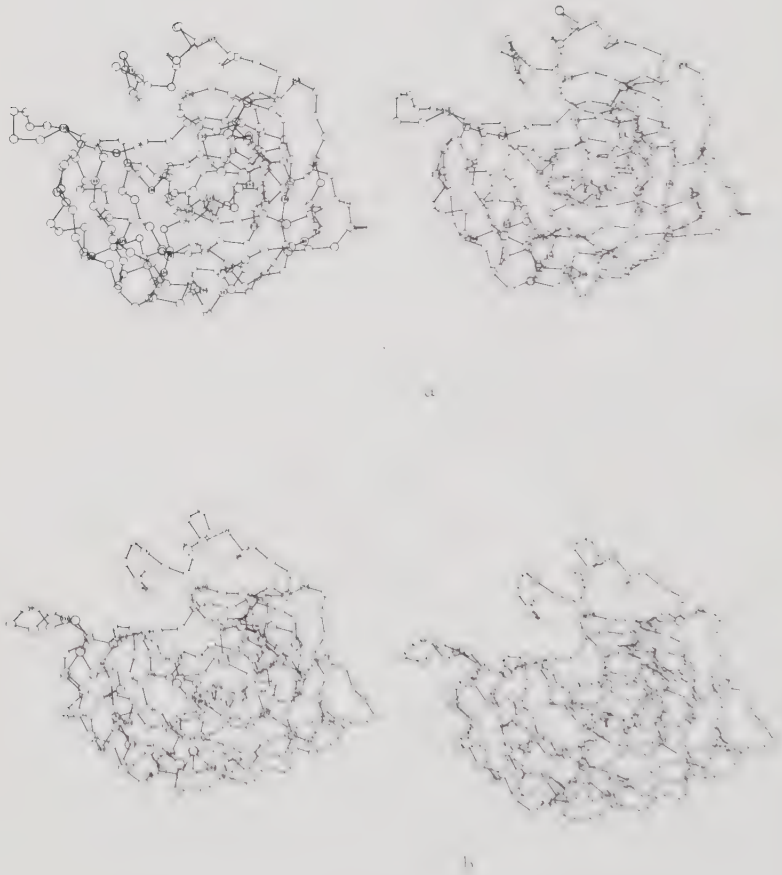


Fig. a A stereoscopic α-carbon diagrams of a) HCAC and b) HCAB. The ORTEP program by Dr. Carrol N. Johnson was was used for the above and subsequent stereodrawings.

be fitted to the electron density. The density is occasionally absent for either the α-carbon or the nitrogen of the peptide bond. Difficulties were encountered in the following three regions.

1) The first three residues at the amino terminus (Ser-His-His) are exposed and do not have any defined density (table 2).

2) Residues 230-237 make a loop out into the solvent without interaction with the rest of the molecule. In this loop there are four glutamyl residues and one aspartyl residue which are exposed and poorly defined in the map (table 2).

3) The backbone region involving the residues 254-256 is poorly defined. However, the side chains except for Ala 256 are well defined and the backbone stretching could be assigned (table 2).

For two proline residues, 29 and 200, which have well defined densities in the map, the cis configuration is the only way to obtain a good fit of the sequence to the electron density. Cis peptides have also been found in other proteins like ribonuclease (37a) and subtilisin (37b).

Recently the cis configuration was observed in synthetic three- and four-cyclic peptides like Pro-Pro-Pro while the more stable transconfiguration was generally more common when the number of peptide units was increased (38). Cis peptide bonds have also been detected in the NMR spectra of co-polypeptides of glycine and proline (39a).

Table 2. The amino acid sequence of HCAC and HCAB, the location of various side chains, secondary structures, polar side chain interactions, aromatic clusters, and the fit of the chemical sequence of HCAC to the electron density. The fit is classified into a scale as follows:

1) Not defined	no density for the side chain
2) Poorly defined	partial density for the side chain
3) Defined	there is density for the whole side chain
4) Well defined	the side chain is defined from the density

β_1, β_2, ... defines the chains in the large pleated sheet that runs through the whole molecule.

β, defines other β-structures.
All interaction distances longer than 3.6 Å are left out.

No	HCAC	HCAB IF DIFFE-RENT	EXPOSED TO SURFACE (S) TO ACTIVE SITE (A)	SECON-DARY STRUC-TURE	POLAR SIDE CHAIN INTER-ACTIONS	ARO-MATIC CLUS-TER	COMMENT
0		ALA					
1	SER						INVISIBLE
2	HIS	PRO					INVISIBLE
3	HIS	ASP				1	INVISIBLE
4	TRP					1	DEFINED
5	GLY		S				WELL DEFINED
6	TYR		A			1	DEFINED
7	GLY	ASP	S	3/10			DEFINED
8	LYS	ASP	S	3/10			NOT DEFINED
9	HIS	LYS	S	3/10			NOT DEFINED
10	ASP	ASN	S	3/10			WELL DEFINED
11	GLY		S				NOT DEFINED
12	PRO	GLN	S	3/10			WELL DEFINED
13	GLU	PRO	S	3/10			DEFINED
14	HIS	GLU	S	3/10			WELL DEFINED
15	TRP			3/10,A		1	WELL DEFINED
16	HIS	SER	S	A			WELL DEFINED
17	LYS		S	A			NOT DEFINED
18	ASP	LEU	S	A			WELL DEFINED
19	PHE	TYR	S	A		1	WELL DEFINED
20	PRO		S	B			WELL DEFINED

NO	HCAC	HCAB IF DIFFE-RENT	EXPOSED TO SURFACE (S) TO ACTIVE SITE (A)	SECON-DARY STRUC-TURE	POLAR SIDE CHAIN INTER-ACTIONS	ARO-MATIC CLUS-TER	COMMENT
21	ILE		S	3/10,B			POORLY DEFINED
22	ALA			3/10,B			WELL DEFINED
23	LYS	ASN	S				NOT DEFINED
24	GLY		S				WELL DEFINED
25	GLU	ASN	S				POORLY DEFINED
26	ARG	ASN	S				POORLY DEFINED
27	GLN				Oε-244Nη		WELL DEFINED
28	SER				Oγ-Trp207Nε \ Ser195NH		WELL DEFINED
29	PRO						DEFINED
30	VAL						DEFINED
31	ASP		S				WELL DEFINED
32	ILE						DEFINED
33	ASP	LYS	S	C			WELL DEFINED
34	THR		S	C			DEFINED
35	HIS	SER	S	C			WELL DEFINED
36	THR	GLU	S	C			WELL DEFINED
37	ALA	THR		C			WELL DEFINED
38	LYS		S	10			NOT DEFINED
39	TYR	HIS	S	10			POORLY DEFINED
40	ASP		S	10	Oδ-Lys255Nη		DEFINED

No	HCAC	HCAB IF DIFFE-RENT	EXPOSED TO SURFACE (S) TO ACTIVE SITE (A')	SECON-DARY STRUC-TURE	POLAR SIDE CHAIN INTER-ACTIONS	ARO-MATIC CLUS-TER	COMMENT
41	PRO	THR	S				NOT DEFINED
42	SER		S				NOT DEFINED
43	LEU						NOT DEFINED
44	LYS		S	β			NOT DEFINED
45	PRO		S	β			WELL DEFINED
46	LEU	ILE		β			POORLY DEFINED
47	SER		S	β			WELL DEFINED
48	VAL			β			WELL DEFINED
49	SER		S	β			WELL DEFINED
50	TYR		S	β			WELL DEFINED
51	ASP	ASN	S	β			NOT DEFINED
52	GLN	PRO	S	β			NOT DEFINED
53	ALA						WELL DEFINED
54	THR		S	β_2			WELL DEFINED
55	SER	ALA		β_2			NOT DEFINED
56	LEU	LYS	A	β_2			WELL DEFINED
57	ARG	GLU	A'	β_2			POORLY DEFINED
58	ILE			β_2			DEFINED
59	LEU	ILE	A'	β_2			WELL DEFINED
60	ASN		A'	β_2			WELL DEFINED

No	HCAC	HCAB IF DIFFE-RENT	EXPOSED TO SURFACE (S) TO ACTIVE SITE (Å)	SECON-DARY STRUC-TURE	POLAR SIDE CHAIN INTER-ACTIONS	AROMATIC CLUS-TER	COMMENT
61	ASN	VAL		β_2			WELL DEFINED
62	GLY						WELL DEFINED
63	HIS		A'		Nδ-Asn66Oδ		WELL DEFINED
64	ALA	SER	A'	β_3			WELL DEFINED
65	PHE			β_3		2	WELL DEFINED
66	ASN	HIS	A'	β_3			WELL DEFINED
67	VAL			β_3			WELL DEFINED
68	GLU	ASN	A'	β_3			WELL DEFINED
69	PHE			β_3		2	WELL DEFINED
70	ASP	GLU	S	β_3			WELL DEFINED
71	ASP		S	β_3			WELL DEFINED
72	SER	ASN	S				WELL DEFINED
73	GLU	ASN	S				WELL DEFINED
74	ASP	ASP	S		O_δ-Arg88Nε ~ Arg88Nη_2		WELL DEFINED
75	LYS	ARG	S	β			POORLY DEFINED
76	ALA	SER		β			WELL DEFINED
77	VAL			β			WELL DEFINED
78	LEU			β			POORLY DEFINED
79	LYS		S	β			NOT DEFINED
80	GLY		S	3/10,β			WELL DEFINED

No	HCAC	HCAB IF DIFFE-RENT	EXPOSED TO SURFACE (S) TO ACTIVE SITE (A)	SECON-DARY STRUC-TURE	POLAR SIDE CHAIN INTER-ACTIONS	ARO-MATIC CLUS-TER	COMMENT
81	GLY		S	3/10, β			WELL DEFINED
82	PRO			3/10			WELL DEFINED
83	LEU	PHE		3/10			WELL DEFINED
84	ASP	SER	S				POORLY DEFINED
85	GLY	ASP	S				WELL DEFINED
86	THR	SER	S	β			WELL DEFINED
87	TYR			β	On-Thr206Oγ		WELL DEFINED
88	ARG		S	β_4			WELL DEFINED
89	LEU			β_4			WELL DEFINED
90	ILE	PHE	A'	β_4			WELL DEFINED
91	GLN		A'	β_4	Oε-His93Nδ		WELL DEFINED
92	PHE			β_4		2	WELL DEFINED
93	HIS		A'	β_4	Nδ-Gln91Oε		WELL DEFINED
94	PHE			β_4		2	WELL DEFINED
95	HIS		A'	β_4			WELL DEFINED
96	TRP					2	WELL DEFINED
97	GLY						WELL DEFINED
98	SER		S		Oγ-Asp241Oδ		WELL DEFINED
99	LEU	THR	S				POORLY DEFINED
100	ASP	ASN	S		Oδ-Arg225Nη		WELL DEFINED

No	HCAC	HCAB IF DIFFERENT	EXPOSED TO SURFACE (S) TO ACTIVE SITE (A')	SECONDARY STRUCTURE	POLAR SIDE CHAIN INTERACTIONS	AROMATIC CLUSTER	COMMENT
101	GLY	GLU	S				NOT DEFINED
102	GLU	HIS	S				POORLY DEFINED
103	GLY						WELL DEFINED
104	SER				Oγ-His106CO		WELL DEFINED
105	GLN	GLU			Oε-Thr197Oγ		WELL DEFINED
106	HIS				CO-Ser104Oγ Nε-Tyr192On Nδ-Glu116Oε1	3	WELL DEFINED
107	THR		S	β			WELL DEFINED
108	VAL			β			WELL DEFINED
109	ASP		S	β			WELL DEFINED
110	LYS	GLY	S				NOT DEFINED
111	LYS						NOT DEFINED
112	LYS	VAL	S				NOT DEFINED
113	TYR		S			7	WELL DEFINED
114	ALA	SER	S				WELL DEFINED
115	ALA			β_5			WELL DEFINED
116	GLU			β_5	Oε1-His106NH Oε2-His106Nδ Oε2-His118Nε		WELL DEFINED
117	LEU			β_5			POORLY DEFINED
118	HIS		A'	β_5	Nε-Glu116Oε2		WELL DEFINED
119	VAL			β_5			POORLY DEFINED
120	VAL	ALA	A'	β_5			WELL DEFINED

No	HCAC	HCAB IF DIFFE-RENT	EXPOSED TO SURFACE (S) TO ACTIVE SITE (A)	SECON-DARY STRUC-TURE	POLAR SIDE CHAIN INTER-ACTIONS	ARO-MATIC CLUS-TER	COMMENT
121	HIS			β_5			WELL DEFINED
122	TRP			β_5			WELL DEFINED
123	ASN			3/10	Nδ-Asp137CO		WELL DEFINED
124	THR	SER	S	3/10			WELL DEFINED
		ALA					
125	LYS		S	3/10			NOT DEFINED
126	TYR		S	3/10			DEFINED
127	GLY		S				DEFINED
128	ASP		S	D	Oδ_1-Lys131Nε		WELL DEFINED
129	PHE	LEU	A′	D			WELL DEFINED
130	GLY		S	D			WELL DEFINED
131	LYS		S	D			WELL DEFINED
132	ALA			D			DEFINED
133	VAL			D			WELL DEFINED
134	GLN		S	D			POORLY DEFINED
135	GLN		S	D			POORLY DEFINED
136	PRO		S				DEFINED
137	ASP		S		Oδ_1-Thr206Oγ Oδ_2-Tyr126Or		WELL DEFINED
138	GLY						WELL DEFINED
139	LEU		A′	β_6			WELL DEFINED
140	ALA			β_6			WELL DEFINED

No	HCAC	HCAB IF DIFFE-RENT	EXPOSED TO SURFACE (S) TO ACTIVE SITE (A)	SECON-DARY STRUC-TURE	POLAR SIDE CHAIN INTER-ACTIONS	ARO-MATIC CLUS-TER	COMMENT
141	VAL		S	β_6			WELL DEFINED
142	LEU	ILE		β_6			WELL DEFINED
143	GLY		S	β_6			WELL DEFINED
144	ILE	VAL		β_6			WELL DEFINED
145	PHE	LEU		β_6		3	WELL DEFINED
146	LEU	MET		β_6			WELL DEFINED
147	LYS		S	β_6			POORLY DEFINED
148	VAL		S	β_6			WELL DEFINED
149	GLY		S	β_6			WELL DEFINED
150	SER	GLU	S				WELL DEFINED
151	ALA		S				WELL DEFINED
152	LYS	ASN	S				DEFINED
153	PRO		S	E			WELL DEFINED
154	GLY	LYS	S	E			WELL DEFINED
155	LEU			E			WELL DEFINED
156	GLN		S	E	$O\varepsilon$-Gln220Nε		WELL DEFINED
157	LYS		S	E			NOT DEFINED
158	VAL			E			WELL DEFINED
159	VAL	LEU		E			POORLY DEFINED
160	ASP		S	E			DEFINED

No	HCAC	HCAB IF DIFFE-RENT	EXPOSED TO SURFACE (S) TO ACTIVE SITE (Å)	SECON-DARY STRUC-TURE	POLAR SIDE CHAIN INTER-ACTIONS	ARO-MATIC CLUS-TER	COMMENT
161	VAL	ALA		E			WELL DEFINED
162	LEU			E			WELL DEFINED
163	ASP	GLN	S	E			POORLY DEFINED
164	SER	ALA	S	E			DEFINED
165	ILE			E			DEFINED
166	LYS		S	E			NOT DEFINED
167	THR		S				WELL DEFINED
168	LYS		S				POORLY DEFINED
169	GLY		S	β_1			WELL DEFINED
170	LYS		S	β_1			NOT DEFINED
171	SER	ARG	S	β_1			WELL DEFINED
172	ALA		S	β_1			WELL DEFINED
173	ASP	PRO	S				POORLY DEFINED
174	PHE						DEFINED
175	THR		S			2	POORLY DEFINED
176	ASN		S				DEFINED
177	PHE					2	WELL DEFINED
178	ASP		S		$O\delta$-Arg180$N\varepsilon$		WELL DEFINED
179	PRO			F			WELL DEFINED
180	ARG	SER	S	F	$N\varepsilon$-Asp178$O\delta$		DEFINED

No	HCAC	HCAB IF DIFFE-RENT	EXPOSED TO SURFACE (S) TO ACTIVE SITE (Å)	SECON-DARY STRUC-TURE	POLAR SIDE CHAIN INTER-ACTIONS	ARO-MATIC CLUS-TER	COMMENT
181	GLY		S	F			WELL DEFINED
182	LEU			F			WELL DEFINED
183	LEU			F			WELL DEFINED
184	PRO						WELL DEFINED
185	GLU	SER	S				POORLY DEFINED
186	SER		S				WELL DEFINED
187	LEU		S				WELL DEFINED
188	ASP		S				WELL DEFINED
189	TYR	PHE		β_8			WELL DEFINED
190	TRP			β_8		3	WELL DEFINED
191	THR		S	β_8			WELL DEFINED
192	TYR			β_8	$O\eta$-Ser280γ	3	WELL DEFINED
193	PRO		S	β_8			WELL DEFINED
194	GLY			β_8			WELL DEFINED
195	SER				Oγ-Leu201CO		WELL DEFINED
196	LEU				NH-Ser280γ		WELL DEFINED
197	THR		A´				WELL DEFINED
198	THR	HIS	A´				DEFINED
199	PRO		A´				WELL DEFINED
200	PRO		A´				WELL DEFINED

No	HCAC	HCAB IF DIFFE-RENT	EXPOSED TO SURFACE (S) TO ACTIVE SITE (A)	SECON-DARY STRUC-TURE	POLAR SIDE CHAIN INTER-ACTIONS	ARO-MATIC CLUS-TER	COMMENT
201	LEU						WELL DEFINED
202	LEU	TYR					WELL DEFINED
203	GLU		S				NOT DEFINED
204	CYS	SER		β_7			WELL DEFINED
205	VAL		A	β_7			DEFINED
206	THR		S	β_7	$O\gamma$-Tyr1260 η -Tyr870 η		WELL DEFINED
207	TRP		A	β_7		3	WELL DEFINED
208	ILE			β_7			WELL DEFINED
209	VAL	ILE	A	β_7			DEFINED
210	LEU	CYS		β_7			WELL DEFINED
211	LYS		S				DEFINED
212	GLU		S				NOT DEFINED
213	PRO	SER	S				WELL DEFINED
214	ILE						WELL DEFINED
215	SER		S				WELL DEFINED
216	VAL						WELL DEFINED
217	SER		S	G			WELL DEFINED
218	SER		S	G			NOT DEFINED
219	GLU		S	G			WELL DEFINED
220	GLN			G			WELL DEFINED

No	HCAC	HCAB IF DIFFE-RENT	EXPOSED TO SURFACE (S) TO ACTIVE SITE (A)	SECON-DARY STRUC-TURE	POLAR SIDE CHAIN INTER-ACTIONS	ARO-MATIC CLUS-TER	COMMENT
221	VAL			G			POORLY DEFINED
222	LEU		S	G			NOT DEFINED
223	LYS		S	G			POORLY DEFINED
224	PHE			G		:	WELL DEFINED
225	ARG			G	Nε-Ser98CO		DEFINED
226	LYS	SER	S	G			NOT DEFINED
227	LEU			G			WELL DEFINED
228	ASN	LEU	S				WELL DEFINED
229	PHE	SER					DEFINED
230	ASP	ASN	S				NOT DEFINED
231	GLY	VAL	S				NOT DEFINED
232	GLU		S				POORLY DEFINED
233	GLY		S				POORLY DEFINED
234	GLU	ASP	S				POORLY DEFINED
235	PRO	ASN	S				POORLY DEFINED
236	GLU	ALA	S				DEFINED
237	GLU	VAL					WELL DEFINED
238	LEU	PRO	S				WELL DEFINED
239	MET						WELL DEFINED
240	VAL	GLN	S				DEFINED

No	HCAC	HCAB IF DIFFE-RENT	EXPOSED TO SURFACE (S) TO ACTIVE SITE (A)	SECON-DARY STRUC-TURE	POLAR SIDE CHAIN INTER-ACTIONS	ARO-MATIC CLUS-TER	COMMENT
241	ASP	HIS	S				WELL DEFINED
242	ASN						WELL DEFINED
243	TRP	ASN					WELL DEFINED
244	ARG						WELL DEFINED
245	PRO		S				WELL DEFINED
246	ALA	THR					WELL DEFINED
247	GLU	GLN	S				WELL DEFINED
248	PRO		S				WELL DEFINED
249	LEU		S				POORLY DEFINED
250	LYS		S				POORLY DEFINED
251	ASN	GLY					POORLY DEFINED
252	ARG				Nε-249CO Nη1-248CO Nη2-193CO		WELL DEFINED
253	GLN	THR					POORLY DEFINED
254	ILE	VAL		β_3			WELL DEFINED
255	LYS	ARG		β_9	Nη-Asp400δ		WELL DEFINED
256	ALA			β_9			NOT DEFINED
257	SER			β_3			WELL DEFINED
258	PHE						DEFINED
259	LYS						WELL DEFINED

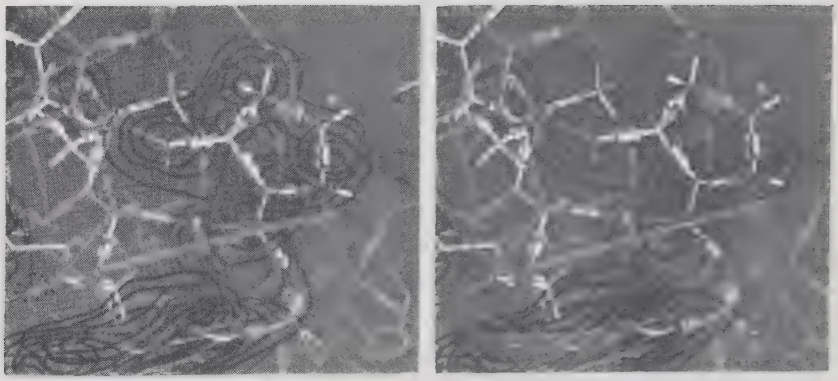


Fig. 6 Stereoscopic photograph of the electron density for

Identification of side chains

Almost all side chains are represented by density in the map except for lysyl residues. The single methionine in HCAC is very well defined with high electron density at the sulfur position. However, the sulfur atom in the only cysteine residue in HCAC shows a very low electron density. As this is the binding site for the methylmercurythioglycolic acid (MMTGA) and the acetoxymercurysulfanilamide (AMSulf) derivatives, the unexpectedly low density may result from imperfect phasing as well as from mobility of the sulfhydryl group.

Aromatic residues are well defined and appear flattened (1). The branched side chains of valine, leucine and isoleucine generally appear forked. Proline residues have well defined densities (except for Pro 41). For example, though it is exposed to solvent, Pro 193 is stabilized in its configuration and has a well defined density (fig. 6). It is hydrogen bonded to the NH group of Arg 252 and in van der Waals contact with His 35 in a neighbouring molecule.

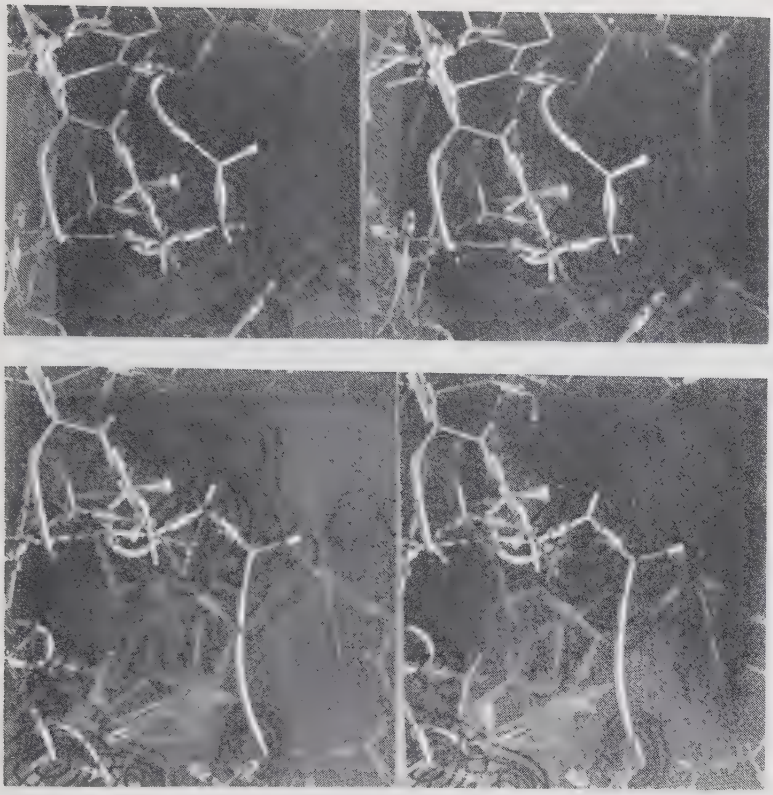


Fig. 7, 8 Stereoscopic photographs of the two possible positions for Asp 137.

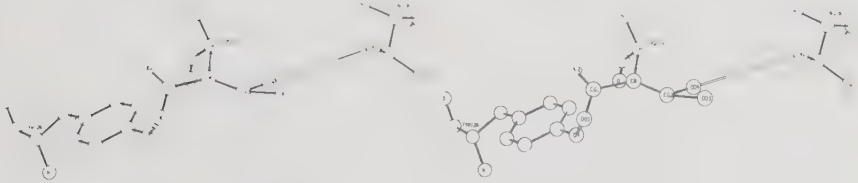


Fig. 9. An ORTEP stereoscopic drawing of the two hydrogen bonding possibilities for Asp 137.

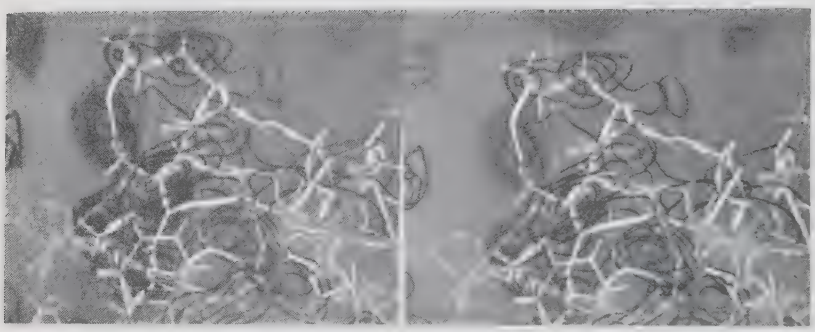


Fig. 10. Stereoscopic photograph of the electron density for Lys 131, which is hydrogenbonded to Asp 128.

The hydrophilic residues are mainly located on the surface of the molecule. The branched part of the hydrophilic side chains give rise to spherical densities, possibly because the carboxylate groups of carboxylic acids, the $-C\begin{smallmatrix}CH_3\\OH\end{smallmatrix}$ of threonines, and the hydroxyl groups of serines are free to rotate. When these groups are involved in hydrogen bonds they are stabilized in their configurations and the densities show the exact positions of the side chain atoms. The side chain of Asp 137 appears to occupy two positions. The electron density indicates that this residue is hydrogen bonded to the hydroxyl group of Tyr 126 in some molecules and in other molecules to the hydroxyl group of threonine 206. Figs 7, 8, 9 show the two possibilities for the location of Asp 137 which are related by rotation around the $C_\alpha - C_\beta$ bond.

Lysyl and arginyl residues are defined only when they are involved in specific polar interactions. Thus only five of the 23 lysyl residues are well defined. It is well known that lysyl residues are flexible or take up two or more fixed positions (39b). Fig. 10 shows the interpretation of Lys 131 which is hydrogen bonded to Asp 128 and is well defined. Three of the five arginyl side chains are well

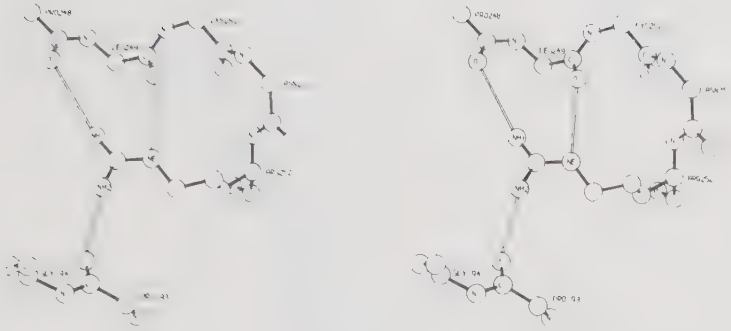


Fig. 11 An ORTEP stereoscopic drawing of Arg 252 and its polar interactions.

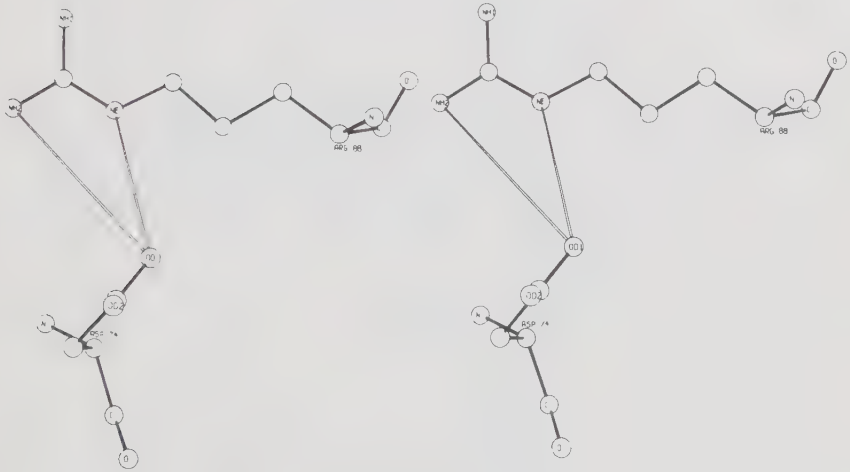


Fig. 12 An ORTEP stereoscopic drawing of Arg 88 and its polar interactions.

defined. The three NH-groups of Arg 252 are hydrogen bonded to the carbonyl groups of Pro 193, Pro 248 and Leu 249 (Fig. 11). The $N\varepsilon$ and the $N\eta$ of Arg 88 are hydrogen bonded to the carboxylate group of Asp 74 (Fig. 12). Residue Arg 225 is stabilized in the same way as Arg 252 except for one NH group, which is exposed to the solvent.

"Model-building"

A device (fig. 13) was built to measure all non-hydrogen atomic coordinates in the model. These coordinates were adjusted to conform to standard bond lengths and angles by a procedure known as "model building", described by Diamond (40). In this procedure the peptide bonds and aromatic rings are assumed to be planar and the side chain atoms are corrected to be in staggered configuration. Pro 29 and 200 were not refined because the "model-building" method does not refine cis peptides. Furthermore because only a limited number of residues could be included in each run the refinement of HCAC was performed in five steps involving the residues 3-28, 30-86, 86-158, 159-199 and 201-259. Therefore, the coordinates of the amino acids in the beginning and the end of each polypeptide segment have a somewhat higher error. The root mean square deviation of all refined coordinates was 0.3 Å. Table 3 lists the atomic coordinates after "model building" refinement and table 4 lists the dihedral angles of all side chains and the χ angles at the C_α-carbons.

The backbone can be described by the dihedral angles (ϕ, ψ), which relate planes of successive peptide bonds to the C_α-atoms linking them (41). A Ramachandran plot (ϕ, ψ) is shown in Fig. 14. The lines in the plot enclose regions outside of which unacceptable van der Waals contacts are formed between C_β-atoms of neighbouring residues. Most of the amino acid residues in the HCAC model except Gly, have dihedral angles in the allowed regions.

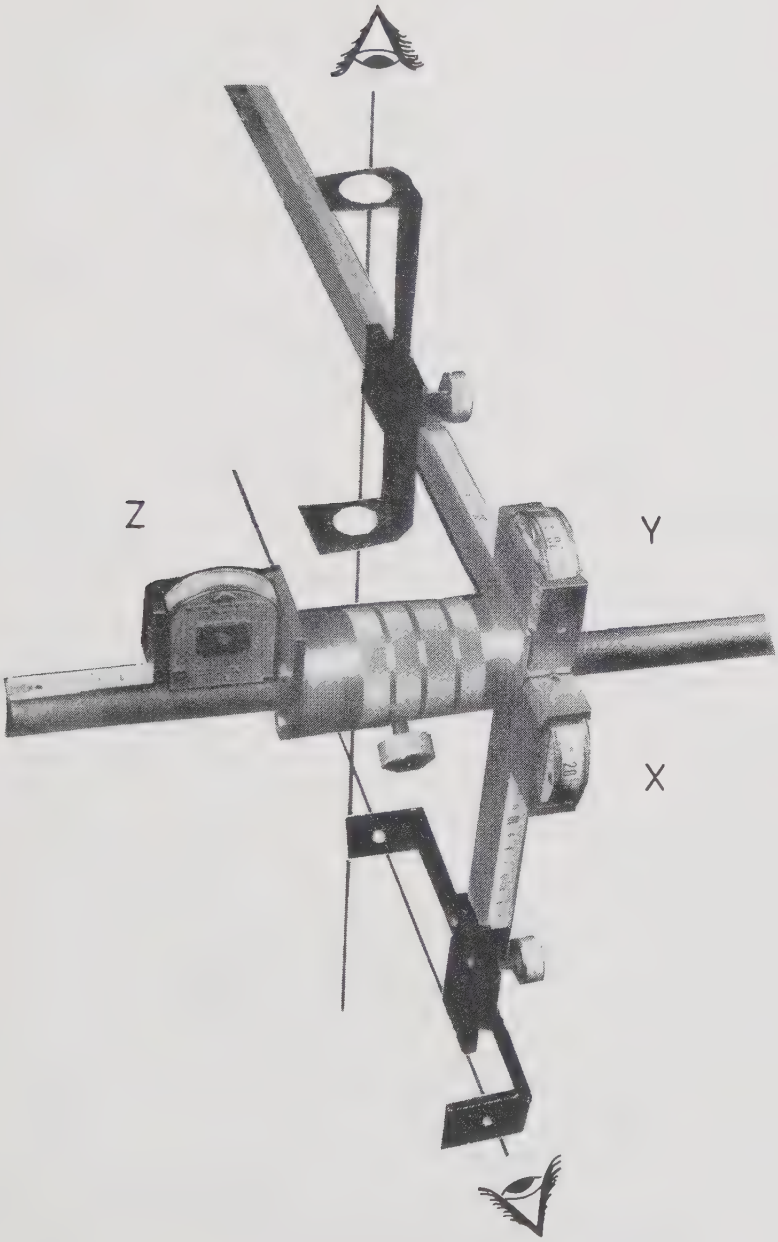
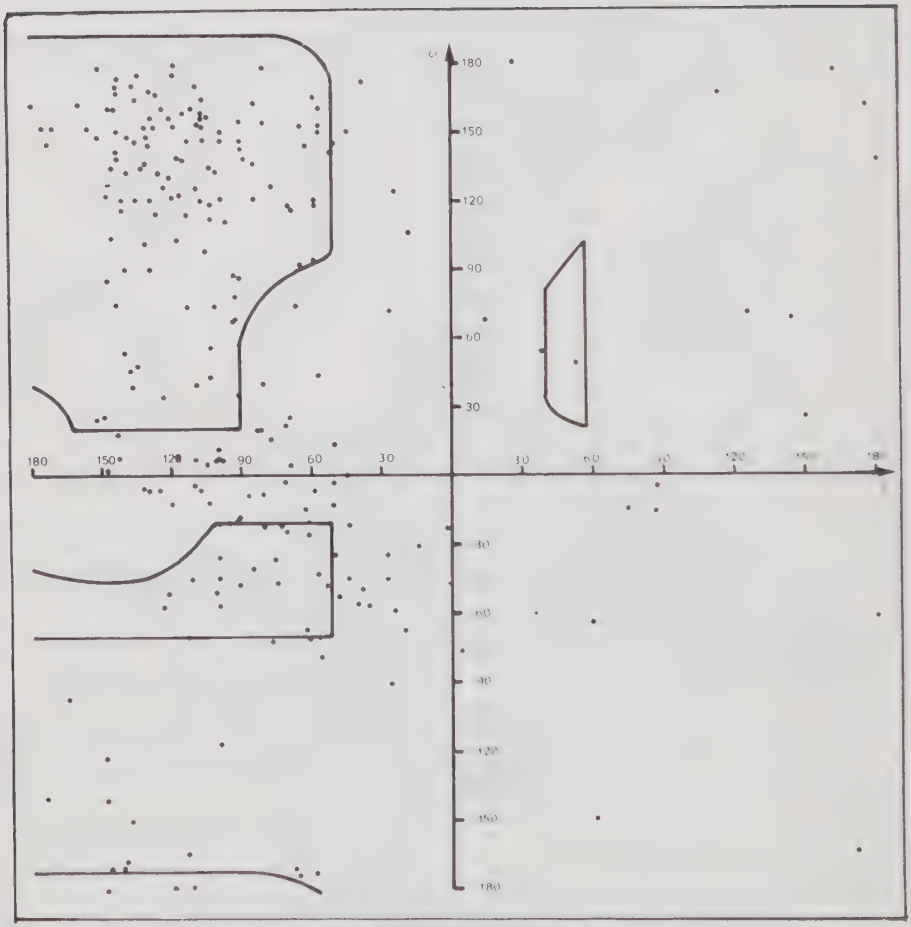


Fig. 13 The measuring device for atomic coordinates.



HCAC RAMACHANDRAN PLOT

Fig. 14

Table 3 Coordinate set for HCAC (scale 2cm=1Å, orthogonal system)

	3 HIS			12 PRO			20 PRO		
CD	9.35	33.41	27.89							
NE	8.01	70.37	29.56	CB	22.16	84.17	43.43	CG	8.17	58.82	52.08
CE	9.38	70.41	31.78	CG	22.63	84.61	46.42	CD	6.57	57.44	54.53
ND	11.63	69.25	31.59	CB	22.27	81.45	47.63	CB	5.80	60.54	55.89
CG	11.07	67.75	29.14	CA	20.29	80.43	45.90	CA	8.02	62.45	55.30
CB	13.45	65.46	28.30	C	17.41	80.04	46.41	C	10.17	62.55	57.48
CA	10.55	67.46	27.65	O	16.00	78.62	47.18	O	10.43	64.55	58.92
C	17.17	69.27	30.00	N	20.52	81.76	43.29	N	9.19	61.45	52.80
O	16.01	71.08	29.91								
N	18.59	65.44	26.89		13 GLU				21 ILE		
				CB	27.70	88.67	54.27				
	4 TRP			CG	26.40	88.36	51.74	CB	19.16	54.51	61.06
CD2	25.00	64.71	37.10	CD	24.11	88.34	52.15	CG1	17.15	56.79	61.83
CE2	24.84	72.43	37.30	OE	22.35	83.33	49.80	CB	15.78	57.91	59.08
CZ2	27.38	73.70	37.57	CB	12.77	86.34	47.56	CG2	17.83	59.21	57.02
CH	29.62	72.57	37.35	CA	13.86	83.57	48.47	CA	13.61	60.22	59.70
CZ1	29.49	69.87	37.41	C	11.57	81.64	47.72	C	15.48	62.78	59.87
CE1	27.63	68.46	37.17	O	9.73	81.22	49.31	O	17.36	62.96	61.48
NE	27.45	69.80	37.02	N	10.59	83.09	47.47	N	11.67	60.39	57.68
CD1	24.04	65.19	36.88								
CG	23.16	67.76	36.83		14 HIS				22 ALA		
CB	20.04	67.75	36.63	CG	10.39	84.26	41.69	CB	15.48	66.59	55.51
CA	19.00	69.60	34.48	ND1	10.09	86.93	40.12	CA	16.21	67.27	58.19
C	21.04	71.68	33.44	CE1	9.04	85.80	38.09	C	15.59	69.33	60.30
O	22.07	71.05	31.60	NE2	9.33	83.54	38.31	O	16.47	71.38	60.61
N	18.31	68.11	32.08	CD2	9.50	82.26	40.58	N	14.76	64.71	58.23
				CB	9.09	79.39	41.39				
	5 GLY			CA	9.65	78.79	44.33		23 LYS		
CA	22.77	76.22	33.70	C	10.21	75.79	44.56	NZ	11.01	80.14	66.86
C	23.82	78.38	35.59	O	8.77	74.14	43.39	CE	11.39	78.14	64.71
O	24.45	77.82	37.43	N	11.72	80.63	45.30	CD	8.84	76.55	64.26
N	20.92	74.08	34.53					CG	8.99	73.80	65.55
					15 TRP			CB	9.89	71.70	63.55
	6 TYR			CD2	20.02	73.74	46.63	CA	12.69	70.60	64.04
OH	33.52	73.81	42.24	CE2	21.99	73.76	49.12	C	12.88	68.77	66.49
CD2	29.43	81.51	40.41	CZ2	24.70	74.52	49.16	O	11.26	68.97	68.36
CE2	30.92	79.70	42.12	CH	26.14	75.32	46.92	N	13.51	68.77	61.89
CZ	32.21	77.80	40.90	CZ1	25.05	75.26	44.42				
CE1	32.41	77.32	38.23	CE1	22.37	74.48	44.36		24 GLY		
CD1	30.99	79.18	36.49	NE1	20.42	74.27	42.27	CA	15.24	65.13	68.67
CG	29.57	81.30	37.04	CD1	18.25	73.30	43.12	C	16.04	66.58	71.23
CB	28.05	83.24	35.84	CG	18.23	73.10	45.77	O	16.00	69.05	71.39
CA	25.30	83.05	36.17	CD	15.85	72.21	47.04	N	14.85	67.02	66.45
C	23.34	85.89	35.70	CA	13.95	72.32	46.42				
O	23.32	87.87	36.26	C	11.69	70.40	47.49		25 GLU		
N	24.00	80.80	34.56	O	12.07	69.70	47.41	OE2	19.63	58.21	76.89
				N	12.29	75.14	46.05	OE1	17.22	57.75	80.55
	7 GLY							CD	17.85	58.79	78.49
CA	20.26	86.64	34.13		16 HIS			CG	16.33	61.32	77.38
C	17.46	88.97	32.94					CB	17.64	64.03	78.06
O	16.70	87.55	31.06	CG	6.63	73.04	52.74	CA	17.75	66.14	75.78
N	21.51	86.04	34.69	ND1	4.43	76.57	53.21	C	20.57	67.30	75.99
				CE1	2.72	75.63	54.69	O	21.15	68.94	77.76
	8 LYS			NE2	3.77	72.88	55.28	N	16.89	65.01	73.20
NZ	17.19	93.56	23.72	CD2	6.25	72.69	54.06				
CE	16.90	95.97	25.34	CB	7.45	70.22	54.28		26 ARG		
CD	14.12	95.37	26.83	CA	8.88	68.92	51.62	NH2	29.90	80.70	86.52
CG	14.44	95.02	29.74	C	6.22	66.05	50.02	NH1	30.28	58.41	82.72
CB	13.22	92.28	30.21	O	4.93	67.72	51.13	CZ	28.95	80.42	83.89
CA	13.24	91.36	33.11	N	10.11	70.87	49.95	NE	27.57	82.21	82.41
C	11.36	88.95	33.29					CD	27.52	82.18	79.47
O	9.36	89.00	34.76		17 LYS			CG	25.79	84.43	78.33
N	15.97	90.84	34.08	NZ	3.03	78.46	40.74	CB	26.93	85.44	75.71
				CE	5.34	74.63	39.52	CA	25.08	87.38	74.18
	9 HIS			CD	4.03	71.76	39.60	C	23.97	87.02	71.27
CD	13.46	83.31	26.12	CG	3.02	70.79	42.47	O	27.44	85.94	70.25
NE	14.16	81.34	24.32	CB	4.77	68.02	42.86	N	22.30	86.43	74.19
CE	12.21	77.64	24.11	CA	4.43	66.88	45.67				
ND	10.22	80.40	25.67	C	4.15	63.84	45.39		27 GLN		
CG	10.97	82.73	26.98	O	1.95	62.74	45.61	NE	21.28	65.80	63.68
CB	9.17	84.16	28.98	N	6.02	62.70	47.44	OE	22.26	68.94	60.65
CA	10.37	84.45	31.76					CD	21.65	66.23	63.00
C	11.40	81.65	32.44		18 ASP			CG	21.17	70.42	65.10
O	9.85	79.72	32.57	OD2	4.15	59.84	41.63	CB	23.56	71.91	65.92
N	12.02	86.89	31.79	OD1	8.16	54.58	40.49	CA	25.09	70.16	67.14
				CG	6.24	55.92	41.15	C	28.46	71.44	66.92
	10 ASP			CB	6.28	59.91	41.57	O	29.77	72.01	66.94
OD2	18.82	79.17	27.38	CA	6.41	59.62	44.57	N	25.00	69.70	69.96
OD1	20.11	80.73	31.03	C	4.67	57.35	46.06				
CG	18.58	79.33	29.76	O	10.45	57.30	45.82		28 SER		
CB	16.41	77.62	30.46	N	6.40	62.54	44.89	OG	33.33	69.38	66.83
CA	15.25	76.90	33.52					CB	34.15	71.15	64.76
C	17.42	79.40	35.62		19 PHE			CA	31.95	73.06	63.96
O	18.98	77.59	36.29	CD2	16.15	63.02	42.86	C	32.76	73.98	61.16
N	14.00	81.47	32.86	CE2	17.01	61.58	40.53	O	31.81	72.89	59.14
				CZ	17.18	58.89	40.52	N	29.32	71.86	64.46
	11 GLY			CE1	16.43	57.49	42.71		29 PRO		
CA	19.47	82.58	38.60	OD1	15.47	58.90	45.00	CD	34.17	77.70	54.67
C	19.01	81.03	41.25	CG	15.41	61.01	45.09	CG	35.80	80.27	57.30
O	17.32	79.22	41.37	CB	14.37	62.97	47.56	CB	37.23	80.57	52.00
N	17.48	81.84	38.63	CA	12.26	61.40	44.08	CA	35.77	78.33	53.00
				C	11.25	62.66	51.69	C	34.23	78.13	55.30
				O	12.34	64.67	52.63	O	32.13	79.23	54.30
				N	9.75	61.41	47.56	N	34.07	76.77	51.09

30 VAL

CG2	37.53	78.43	74.22
CG1	32.54	77.77	73.59
CB	34.98	79.63	73.17
CA	35.33	80.75	70.33
C	37.06	83.20	70.59
O	36.99	83.56	69.10
N	33.33	69.38	66.83

31 ASP

OD2	39.13	92.95	72.72
OE1	36.67	93.53	76.05
CG	37.25	92.27	74.05
CB	35.86	89.76	73.12
CA	37.82	87.39	72.93
C	39.54	87.24	75.46
O	38.92	85.84	77.37
N	36.32	84.90	72.48

32 ILE

CD	50.87	86.26	77.27
CG1	47.80	90.15	76.89
CB	46.47	88.98	76.63
CG2	47.66	91.08	78.72
CA	43.51	88.93	77.73
C	43.55	91.16	79.83
O	44.07	93.50	79.19
N	41.68	88.84	75.43

33 ASP

OD2	42.14	95.20	88.84
OD1	37.99	94.97	89.27
CG	40.03	94.32	88.11
CB	40.17	92.33	85.85
CA	42.94	92.35	84.50
C	45.30	91.03	85.93
O	45.00	88.92	87.20
N	42.99	90.41	82.30

34 THR

CG	52.03	93.05	82.68
OG	52.70	95.07	87.01
CB	52.45	92.59	85.63
CA	50.04	91.62	86.97
C	49.90	91.72	84.98
O	50.81	90.97	91.59
N	47.62	92.28	85.70

35 HIS

CD	50.40	101.19	92.63
NE	50.56	102.95	90.57
CE	48.61	102.51	88.87
ND	47.11	100.55	89.79
CG	48.20	99.66	92.18
CH	47.00	97.37	93.74
CA	46.5[illegible]	94.76	93.53
C	47.11	92.62	95.23
O	48.09	91.89	97.37
N	48.77	94.00	90.70

36 THR

CG	38.49	88.00	95.77
OG	39.79	[illegible]	97.24
CB	40.29	90.37	95.30
CA	43.27	93.68	[illegible]
C	43.96	[illegible]	96.67
O	43.48	[illegible]	[illegible]
N	44.84	91.71	94.25

37 ALA

CB	47.78	[illegible]	[illegible]
CA	45.82	[illegible]	91.07
C	47.39	[illegible]	[illegible]
O	48.41	[illegible]	94.86
N	45.07	86.79	92.28

38 LYS

NZ	41.43	80.88	103.31
CE	42.14	81.44	100.90
CD	43.87	77.98	[illegible]
CG	43.99	77.03	97.43
CB	46.84	[illegible]	96.52
CA	46.50	79.71	95.59
C	[illegible]	77.15	96.29
O	51.30	76.67	[illegible]
N	47.11	80.26	93.50

39 TYR

CD2	50.14	85.34	105.45
CE2	55.26	84.27	106.31
CZ	53.54	85.18	107.52
CE1	51.85	83.82	106.31
CD1	51.49	[illegible]	[illegible]
CG	53.71	79.55	103.00
CB	55.42	80.88	101.19
CA	57.27	79.17	99.44
C	55.73	77.33	97.52
O	56.75	76.72	98.73
N	57.64	76.66	101.32
N	53.17	78.61	96.82

40 ASP

OD2	58.57	84.71	98.99
OD1	56.42	83.13	95.77
CG	56.87	83.33	97.24
CB	55.63	81.76	96.74
CA	57.45	84.96	[illegible]
C	60.30	84.31	97.59
O	60.82	87.56	97.04
N	56.55	72.59	97.18

41 PRO

CD	61.58	71.14	101.88
CG	64.16	70.71	103.47
CB	66.30	70.33	101.34
CA	64.95	84.09	99.31
C	65.23	86.34	98.99
O	[illegible]	89.85	97.30
N	62.12	69.84	99.31

42 SER

OG	61.69	81.58	105.36
CB	62.30	84.80	102.64
CA	64.56	81.90	[illegible]
C	65.31	80.07	99.03
O	66.21	[illegible]	94.27
N	64.41	84.82	101.19

43 LEU

CD1	58.72	[illegible]	91.53
CG	61.30	81.19	87.45
CD2	61.39	82.30	90.40
CB	61.75	83.36	92.59
CA	64.33	83.39	94.27
C	65.56	80.21	92.44
O	67.49	82.30	[illegible]
N	64.06	80.96	96.74

44 LYS

CD	71.23	93.37	93.37
CE	73.83	[illegible]	[illegible]
NZ	73.74	92.39	[illegible]
CG	71.66	[illegible]	90.80
CH	71.55	[illegible]	88.66
CA	69.92	88.78	84.76
C	[illegible]	[illegible]	86.35
O	67.49	[illegible]	[illegible]
N	67.68	88.26	90.94

45 PRO

CD	73.41	82.61	[illegible]
CG	74.43	84.04	84.80
CB	76.76	[illegible]	[illegible]
CA	76.58	[illegible]	82.67
C	[illegible]	[illegible]	80.72
O	78.56	84.63	81.40
N	74.41	81.78	85.52

46 LEU

CD1	[illegible]	[illegible]	[illegible]
CG	81.12	85.32	71.79
CD2	[illegible]	[illegible]	[illegible]
CB	[illegible]	[illegible]	73.50
CA	71.41	[illegible]	[illegible]
C	74.09	[illegible]	74.96
O	76.13	[illegible]	75.17
N	71.20	[illegible]	[illegible]

47 [illegible]

CD	77.99	90.63	[illegible]
CB	76.81	[illegible]	71.80
CA	[illegible]	93.33	[illegible]
C	[illegible]	[illegible]	[illegible]
O	73.74	94.39	[illegible]
N	73.96	96.55	73.97

40 VAL

CG2	75.27	[illegible]	65.84
CG1	76.49	94.96	61.96
CB	77.15	94.23	64.27
CA	77.52	96.35	65.00
C	79.04	[illegible]	63.39
O	[illegible]	96.97	62.70
N	77.78	96.27	67.92

44 SER

OG	81.70	[illegible]	[illegible]
CB	81.69	98.23	[illegible]
CA	81.72	[illegible]	61.30
C	80.54	96.15	[illegible]
O	78.29	[illegible]	57.91
N	79.62	92.80	62.82

50 TYR

CD2	70.14	90.94	56.23
CE2	70.15	89.66	55.55
CZ	73.35	[illegible]	55.91
CE1	72.51	90.14	55.89
CD1	74.65	92.06	55.40
CG	77.45	91.23	55.04
CB	78.37	93.50	55.06
CA	81.00	97.66	54.70
C	81.37	99.03	54.52
O	82.57	[illegible]	51.88
N	81.87	99.08	[illegible]
N	82.21	93.61	57.14

51 ASP

OD2	92.23	90.59	52.71
OD1	90.03	93.42	53.06
CG	90.77	92.96	52.87
CB	89.83	93.58	50.25
CA	88.41	92.75	49.66
C	87.37	[illegible]	48.99
O	89.25	89.31	[illegible]
N	85.25	93.17	52.05

52 GLN

NE2	[illegible]	88.31	[illegible]
CD	88.76	90.80	53.57
CG	88.69	[illegible]	51.07
CB	[illegible]	[illegible]	[illegible]
CA	[illegible]	93.98	49.32
C	[illegible]	95.96	[illegible]
O	81.93	[illegible]	[illegible]
N	80.27	[illegible]	48.70
N	84.78	98.66	48.23

53 ALA

CB	77.71	92.82	[illegible]
CA	78.59	[illegible]	[illegible]
C	79.23	[illegible]	41.66
O	[illegible]	[illegible]	40.85
N	[illegible]	98.97	45.82

54 THR

CG	78.89	[illegible]	31.66
OG	[illegible]	[illegible]	38.26
CB	[illegible]	95.98	[illegible]
CA	77.67	98.57	37.17
C	[illegible]	[illegible]	35.25
O	[illegible]	[illegible]	33.11
N	77.27	98.79	40.08

55 SER

OG	70.65	97.64	33.21
CB	71.03	[illegible]	33.74
CA	72.50	93.23	34.13
C	[illegible]	[illegible]	[illegible]
O	[illegible]	92.90	34.42
N	74.57	92.30	35.82

56 LEU

CD1	65.23	93.84	30.83
CG	64.52	[illegible]	28.29
CD2	66.60	93.22	[illegible]
CB	66.37	[illegible]	[illegible]
CA	[illegible]	93.90	28.30
C	[illegible]	[illegible]	25.96
O	[illegible]	91.88	24.02
N	64.71	92.21	30.95

57 ARG

NH2	60.81	40.53	21.81
NH1	64.75	44.16	22.22
CZ	63.45	40.40	21.84
NE	64.81	43.74	21.52
CD	55.05	30.17	18.93
CG	54.55	33.15	19.23
CB	63.90	33.75	22.03
CA	65.41	55.05	23.97
C	64.15	57.55	25.10
O	63.46	57.72	27.48
N	60.19	53.19	20.11

58 ILE

CD	60.67	60.37	26.29
CG1	63.69	67.00	25.63
CB	64.84	64.34	24.52
CG2	66.02	64.40	21.70
CA	62.69	62.14	24.22
C	60.70	62.74	22.04
O	61.11	62.34	19.80
N	63.87	59.58	23.38

59 LEU

CD2	52.72	62.40	24.14
CD1	53.33	59.35	21.50
CG	52.85	61.33	21.35
CB	55.07	62.63	19.70
CA	56.48	65.09	20.98
C	54.08	66.88	21.60
O	54.30	66.32	23.13
N	58.56	64.23	22.87

60 ASN

ND2	45.96	73.16	16.75
OD	44.72	70.35	20.41
CG	46.43	71.26	18.72
CB	49.13	69.96	18.66
CA	49.37	67.76	20.83
C	46.08	66.30	20.85
O	45.98	64.90	18.93
N	51.82	66.20	20.42

61 ASP

OD2	40.13	66.77	26.73
OD1	37.98	63.30	27.79
CG	39.99	64.67	27.57
CB	42.46	63.87	26.03
CA	42.01	65.33	23.31
C	40.00	66.92	23.16
O	37.63	65.87	23.72
N	45.23	66.63	23.03

62 GLY

CA	37.90	71.19	22.18
C	36.93	72.65	24.69
O	35.65	74.26	24.58
N	40.25	69.44	22.40

63 HIS

CD	36.30	67.29	33.82
NE	37.74	66.89	36.09
CE	38.91	64.13	36.80
ND	38.22	71.03	35.10
CG	36.56	69.92	33.18
CB	35.46	71.24	30.99
CA	37.53	71.31	29.47
C	40.10	73.64	31.04
O	40.04	74.56	33.34
N	38.23	72.06	26.91

64 ALA

CB	44.77	71.38	33.61
CA	44.93	73.08	31.06
C	47.32	72.03	29.48
O	47.01	70.47	27.58
N	42.29	72.88	29.77

65 PHE

CD2	58.27	70.17	30.53
CE2	60.97	75.66	31.37
CZ	62.02	73.20	31.33
CE1	60.56	71.13	30.36
CD1	57.93	71.62	29.45
CG	56.77	74.07	29.60
CB	53.97	74.46	28.57
CA	52.15	72.06	28.92
C	53.07	69.92	30.91
O	53.11	70.38	33.34
N	49.68	72.90	30.28

66 ASN

ND	51.65	59.53	34.66
OD	51.40	63.79	36.12
CG	51.89	62.67	34.28
CB	52.62	63.13	31.49
CA	54.72	65.40	31.52
C	57.41	64.14	30.80
O	58.31	64.29	28.49
N	53.80	67.62	29.84

67 VAL

CG2	66.17	62.75	32.14
CG1	63.82	63.08	30.39
CB	63.49	63.47	33.32
CA	61.27	61.64	32.36
C	61.22	58.99	33.88
O	59.93	58.76	35.98
N	58.66	62.92	32.78

68 GLU

OE2	55.00	53.92	28.78
OE1	58.44	51.59	27.36
CD	57.15	52.76	29.00
CG	58.38	52.99	32.00
CB	61.05	52.29	32.49
CA	62.73	54.38	34.06
C	65.47	53.05	34.37
O	67.09	53.01	32.49
N	62.63	57.03	32.80

69 PHE

CD2	73.15	55.36	40.06
CE2	73.81	56.15	40.05
CZ	71.93	60.04	39.68
CE1	69.33	59.32	39.44
CD1	68.64	56.55	39.53
CG	70.52	54.62	39.77
CB	69.66	51.75	39.89
CA	68.51	50.66	37.30
C	67.65	47.67	37.42
O	65.50	46.84	37.45
N	65.95	51.99	36.74

70 ASP

OD2	73.03	37.67	35.27
OD1	70.74	37.41	36.76
CG	72.01	38.74	37.17
CB	72.32	41.74	37.28
CA	69.58	43.14	37.58
C	68.24	42.15	40.15
O	66.47	43.41	42.27
N	69.93	46.05	37.48

71 ASP

OD2	61.75	40.77	45.61
OD1	58.73	41.48	42.81
CG	60.85	40.43	43.41
CB	62.43	38.59	41.63
CA	65.46	38.76	42.22
C	65.51	36.20	43.90
O	64.29	36.07	46.06
N	66.85	39.92	39.91

72 SER

OG	66.83	30.23	39.65
CB	67.31	29.36	42.32
CA	67.13	31.65	44.31
C	69.42	31.09	46.26
O	71.18	29.43	45.72
N	66.92	34.22	42.90

73 GLU

OE2	75.39	27.17	53.36
OE1	74.12	25.84	49.27
CD	74.99	27.39	50.91
CG	76.30	30.19	50.13
CB	74.30	32.35	49.31
CA	71.45	32.17	50.54
C	71.11	34.37	52.63
O	68.97	35.61	52.84
N	69.36	32.50	48.49

74 ASP

OD2	70.07	35.92	62.67
OD1	68.33	37.39	59.16
CG	70.23	36.40	60.32
CB	72.81	35.56	59.01
CA	73.16	36.90	56.24
C	75.65	38.66	56.29
O	76.66	39.39	58.44
N	73.22	34.84	54.15

75 LYS

NZ	81.28	35.83	43.65
CE	79.38	37.78	44.35
CD	78.82	38.00	47.34
CG	79.81	40.57	48.65
CB	80.85	40.03	51.46
CA	79.02	41.01	53.70
C	78.78	44.01	53.15
O	80.79	45.38	52.65
N	76.62	39.33	53.93

76 ALA

CB	74.31	48.60	50.06
CA	75.82	47.83	52.74
C	73.94	47.69	54.78
O	71.21	47.27	54.06
N	76.33	44.98	53.24

77 VAL

CG2	74.95	45.23	62.75
CG1	72.30	42.76	59.20
CB	72.52	45.29	60.97
CA	72.27	47.93	59.46
C	72.10	50.01	61.71
O	73.78	51.82	61.89
N	74.26	48.04	57.30

78 LEU

CD2	62.94	54.51	63.56
CG1	64.44	55.21	68.22
CG	64.83	53.57	65.72
CB	67.71	53.81	64.80
CA	69.68	51.58	65.61
C	68.61	50.44	68.24
O	66.71	48.86	68.27
N	70.08	49.71	63.37

79 LYS

NZ	77.42	41.08	73.76
CE	74.98	42.33	72.65
CD	74.97	45.33	73.17
CG	72.46	46.65	72.02
CB	70.99	48.22	74.16
CA	69.05	50.32	73.08
C	69.06	52.95	74.64
O	69.98	55.02	73.64
N	69.87	51.27	70.42

80 GLY

CA	67.91	55.21	78.79
C	65.73	55.26	80.90
O	63.57	54.05	80.54
N	68.04	52.82	77.08

81 GLY

CA	64.29	56.86	85.27
C	61.35	56.94	84.43
O	59.73	55.29	85.31
N	66.24	56.66	83.07

82 PRO

CD	62.58	60.55	81.27
CG	60.94	61.64	78.94
CB	58.07	61.53	79.94
CA	57.94	59.18	81.79
C	57.02	56.60	80.44
O	54.90	56.51	79.17
N	60.71	58.87	82.75

83 LEU

CD2	55.05	52.74	74.40
CD1	59.26	52.31	71.86
CG	58.03	53.21	74.46
CB	59.29	51.69	76.76
CA	57.99	51.93	79.54
C	58.76	49.34	80.98
O	60.52	49.32	82.72
N	58.63	54.53	80.75

84 ASP

OD2	52.11	46.55	83.81
OD1	51.82	43.01	81.60
CG	53.14	44.56	82.95
CB	56.00	44.10	83.80
CA	57.95	44.57	81.46
C	57.61	42.63	79.12
O	55.35	42.02	78.29
N	57.42	47.19	80.23

85 GLY

CA	59.80	39.80	75.83
C	60.90	40.95	73.22
O	61.25	43.39	72.95
N	59.85	41.68	78.09

86 THR

CG	64.42	37.45	64.52
CB	66.13	37.07	68.93
CB	63.76	37.52	67.42
CA	62.50	43.01	66.74
C	60.64	41.30	66.68
O	58.55	40.17	65.97
N	61.42	39.17	71.34

87 TYR

CH	49.81	47.64	71.62
CD2	56.63	46.40	64.14
CE2	54.30	46.29	70.83
CZ	52.03	47.70	69.98
CE1	51.89	43.44	67.62
CD1	54.25	43.04	65.91
CG	56.55	42.70	66.73
CB	59.04	42.83	64.93
CA	59.83	45.16	63.76
C	61.52	45.32	61.23
O	63.44	45.85	61.35
N	61.43	43.64	65.74

88 ARG

NH2	64.72	30.42	52.94
NH1	60.20	31.32	52.96
CZ	62.74	32.15	53.08
NE	63.32	34.78	53.34
CD	61.28	36.90	53.30
CG	62.40	39.51	54.36
CB	60.11	41.36	55.26
CA	61.03	44.02	56.51
C	60.15	46.33	54.71
O	57.74	46.90	54.91
N	59.94	44.37	59.22

89 LEU

CD2	67.48	53.33	48.82
CD1	63.67	56.38	49.61
CG	64.51	53.49	49.35
CB	63.81	52.00	51.89
CA	61.52	49.94	51.76
C	61.14	48.76	48.90
O	62.61	47.21	47.95
N	62.10	47.67	53.55

90 ILE

CG	52.25	40.01	41.33
CD1	55.13	47.09	41.66
CB	55.67	48.36	44.47
CG2	54.07	50.85	45.17
CA	56.59	47.14	44.79
C	54.84	51.42	42.57
O	60.23	53.98	40.56
N	59.18	50.14	47.57

91 GLN

NE	49.78	57.62	41.34
CD	53.13	60.39	42.96
CG	52.03	58.73	41.37
CB	53.57	57.90	38.84
CA	55.20	55.47	39.12
C	57.55	55.32	40.46
O	57.49	58.20	41.38
N	56.84	57.91	45.26
N	57.04	53.74	43.01

92 PHE

CD2	65.73	60.62	41.16
CE2	67.65	64.90	41.48
CZ	66.44	66.78	39.69
CE1	64.16	67.00	39.55
CD1	62.43	66.90	41.27
CG	63.08	64.41	42.99
CB	60.45	64.26	44.77
CA	58.23	62.47	41.47
C	57.57	60.17	41.70
O	58.16	59.36	39.28
N	58.19	60.61	41.44

93 HIS

CG	48.26	68.19	42.46
NE	47.47	66.51	44.25
CE	49.53	64.96	44.86
NE	51.64	65.82	43.57
CD	50.91	67.01	42.03
CB	52.88	66.45	41.03
CA	55.66	64.84	41.40
C	55.52	72.61	42.74
O	56.75	73.11	44.84
N	56.40	67.48	42.47

94 PHE

CD2	60.43	73.46	42.89
CE2	63.16	78.57	42.97
CZ	63.44	76.32	41.70
CE1	62.15	74.13	40.21
CD1	59.46	75.83	40.35
CG	58.61	78.03	41.41
CB	55.75	78.92	41.13
CA	53.73	77.12	42.49
C	51.27	78.52	42.22
O	49.45	77.71	40.39
N	54.01	74.41	41.47

95 HIS

CG	46.13	75.83	46.49
NE	44.38	74.07	46.23
CE	41.52	75.38	45.47
ND	42.34	77.97	45.96
CG	45.37	78.32	46.30
CB	46.33	81.06	46.41
CA	47.47	81.85	44.30
C	48.15	84.90	43.97
O	49.59	86.41	45.61
N	50.49	80.39	44.07

96 TRY

CD2	48.91	83.90	36.38
CE2	48.81	83.68	37.92
CZ2	48.71	81.27	38.47
CH	48.82	81.10	33.70
CZ1	48.83	83.32	34.44
CE1	48.91	83.73	33.54
NE	48.95	86.37	32.52
CD1	49.13	87.97	34.70
CG	49.01	86.68	37.05
CD	49.04	87.74	39.47
CA	46.73	89.34	41.83
C	44.49	90.74	40.85
O	42.13	89.59	40.68
N	46.77	86.12	42.39

97 GLY

CA	41.11	93.11	39.28
C	43.63	93.87	38.79
O	45.25	96.13	35.06
N	45.13	93.28	40.23

98 SER

OG	38.95	104.02	35.39
CB	39.90	102.48	34.53
CA	42.57	102.41	34.14
C	44.47	103.27	34.85
O	45.41	104.17	33.04
N	42.26	99.12	36.04

99 LEU

CD2	32.44	111.11	37.01
CD1	34.96	109.02	37.13
CG	42.04	110.77	37.12
CB	44.47	110.41	39.13
CA	46.16	106.99	38.07
C	44.42	105.49	40.11
O	47.24	103.19	41.12
N	44.96	104.23	37.31

100 ASP

CG	47.45	106.21	41.44
OD1	46.13	106.76	38.87
CB	46.38	107.21	40.72
CG	44.11	108.62	43.04
CA	51.44	106.53	43.33
C	50.53	106.38	45.74
O	51.48	105.05	47.01
N	49.94	107.24	43.49

101 GLY

CA	46.73	102.77	48.44
C	44.90	103.39	48.95
O	44.11	105.55	50.54
N	48.09	107.75	45.92

102 GLU

OE2	36.11	111.11	44.27
CD	37.14	101.36	41.42
OE1	37.77	100.14	39.16
CG	40.11	103.20	40.16
CB	40.36	104.16	37.47
CA	43.44	103.74	37.41
C	44.41	104.41	36.04
O	45.77	98.76	43.46
N	45.49	103.20	47.60

103 GLY

CA	46.04	93.64	45.25
C	44.31	91.44	47.33
O	43.23	84.20	46.73
N	43.62	90.06	46.85

104 SER

OG	45.85	93.44	36.28
CB	45.54	91.44	34.80
CA	43.53	90.06	31.92
C	41.54	87.79	32.64
O	39.28	87.76	31.75
N	43.42	92.01	49.74

105 GLN

NE2	38.21	79.20	49.94
CD	41.18	79.84	50.63
OE1	40.37	78.29	50.70
CG	42.36	80.48	51.52
CB	42.40	81.11	54.39
CA	40.92	83.04	55.37
C	40.50	84.30	58.04
O	38.42	85.04	58.96
N	42.56	85.92	54.24

106 HIS

CG	44.30	77.73	62.23
ND1	43.11	75.56	63.34
CE1	42.06	76.22	65.66
NE2	41.49	78.79	66.06
CD2	43.89	79.82	63.92
CB	44.74	82.71	63.78
CA	42.75	84.57	62.39
C	43.46	87.53	62.68
O	44.83	88.69	60.99
N	42.78	84.01	59.51

107 THR

OG1	39.98	89.28	62.25
CB	38.49	91.12	65.75
CG2	40.41	93.05	65.06
CA	39.28	91.55	65.38
C	44.28	91.65	68.19
O	44.22	89.66	69.67
N	42.55	88.70	64.86

108 VAL

CG2	30.72	94.44	64.65
CG1	31.15	96.42	73.93
CB	33.01	94.67	71.05
CA	30.98	94.32	71.53
C	29.44	96.87	72.78
O	28.25	97.11	72.10
N	45.35	93.96	68.89

109 ASP

OD2	45.11	96.09	80.03
CG	45.27	99.37	82.54
OD1	44.74	98.47	80.32
CB	44.29	100.13	77.84
CA	43.03	98.78	75.94
C	41.18	100.88	74.31
O	41.68	103.29	74.42
N	43.55	96.50	74.59

110 LYS

NZ	29.15	100.38	72.74
CE	30.98	101.66	71.02
CD	33.31	103.16	72.70
CG	34.72	105.33	70.99
CB	37.67	104.39	71.28
CA	38.53	101.75	74.03
C	39.95	102.75	67.52
O	38.68	103.80	65.67
N	39.70	99.94	72.04

111 LYS

NZ	41.25	107.92	63.16
CE	42.34	104.76	65.91
CD	43.41	104.59	65.97
CG	44.08	104.71	67.74
CB	45.55	104.93	66.13
CA	44.19	103.29	65.20
C	45.13	101.01	63.41
O	45.74	98.80	64.37
N	42.57	102.41	67.49

112 LYS

NZ	37.45	98.93	51.87
CE	37.05	99.43	59.10
CD	34.38	100.24	57.33
CG	41.54	98.14	57.21
CB	44.24	99.38	56.53
CA	46.17	99.59	58.07
C	46.87	97.03	57.43
O	50.33	101.78	57.03
N	45.28	101.57	60.83

113 TYR

OH	58.17	97.72	66.74
CZ	57.73	98.33	59.51
CE2	58.80	97.26	62.05
CZ	57.15	97.00	64.28
CE1	54.80	95.95	64.14
CD1	53.50	95.27	61.53
CG	55.21	95.53	59.27
CB	54.32	94.79	56.53
CA	52.39	97.09	55.28
C	51.97	98.04	52.29
O	49.87	97.03	51.93
N	49.81	97.29	56.67

114 ALA

CB	58.69	97.08	46.72
CA	53.82	97.04	47.82
C	52.55	94.45	46.92
O	50.57	94.43	45.44
N	53.97	97.40	50.74

115 ALA

CB	54.21	88.55	44.64
CA	52.72	89.80	47.12
C	53.36	87.67	49.39
O	54.65	88.41	51.37
N	53.72	92.27	47.83

116 GLU

OE2	47.85	78.58	57.71
OE1	46.53	82.78	57.72
CD	48.02	80.99	57.20
CG	50.70	81.46	55.81
CB	50.33	82.49	52.72
CA	52.32	83.15	51.13
C	53.92	80.61	49.76
O	52.54	79.45	48.06
N	52.45	85.21	49.08

117 LEU

CD2	62.07	74.00	46.29
CD1	61.12	72.96	51.03
CG	62.23	73.24	49.12
CB	60.59	77.59	49.39
CA	57.52	77.37	49.35
C	56.63	74.88	50.89
O	56.80	74.07	53.37
N	56.31	79.79	50.52

118 HIS

CD	50.17	71.35	54.20
NE	47.62	71.87	55.08
CE	45.96	71.42	53.10
ND	47.34	70.58	51.01
CG	50.03	70.54	51.09
CB	52.16	69.74	49.72
CA	55.22	72.32	50.63
C	56.41	67.80	50.31
O	56.44	66.44	48.12
N	55.93	72.82	49.40

119 LEU

CD2	64.64	67.91	55.03
CD1	63.20	66.82	50.43
CG	63.85	65.88	53.15
CB	61.45	64.19	54.23
CA	58.93	64.06	52.46
C	57.27	61.67	53.42
O	56.77	61.35	55.62
N	57.52	66.64	52.50

120 VAL

CG2	51.43	59.87	49.15
CG1	50.58	55.09	50.65
CB	52.69	57.27	50.91
CA	54.81	57.68	52.16
C	56.12	54.94	52.31
O	57.44	53.92	50.58
N	56.43	60.04	51.52

121 HIS

CD	59.04	56.51	58.62
NE	61.65	56.37	58.11
CE	63.03	57.12	56.91
NE	62.95	58.60	56.56
CG	60.43	58.21	57.62
CB	59.08	57.57	57.60
CA	58.92	57.18	59.47
C	59.36	59.74	58.32
O	58.28	61.02	58.79
N	55.75	53.80	54.86

122 TRY

CZ2	52.80	40.79	51.91
CE2	52.15	38.65	53.30
CZ	53.39	38.25	52.17
OH	54.07	38.11	49.89
CZ1	55.30	38.33	48.44
CE1	54.35	40.72	49.53
NE	54.07	41.05	48.54
CD1	53.45	44.95	50.36
CG	52.23	43.64	52.36
CB	50.81	44.79	54.80
CA	52.22	45.50	57.30
C	53.45	43.12	58.76
O	55.67	42.22	58.10
N	54.54	47.12	56.51

123 ASN

ND	49.93	33.02	64.23
OD	54.09	40.54	67.75
CG	51.84	39.78	67.21
CB	50.87	37.04	64.45
CA	52.95	39.83	62.30
C	53.73	37.05	61.31
O	52.10	35.58	60.20
N	52.00	42.12	60.73

124 THR

CG	60.18	32.78	64.91
OG	61.30	31.37	60.44
CB	60.16	33.48	61.98
CA	57.30	33.77	60.86
C	55.84	31.17	61.53
O	55.68	29.24	59.97
N	56.27	36.38	61.72

125 LYS

NZ	62.06	30.26	69.17
CE	60.32	28.06	70.10
CD	57.43	28.98	70.32
CG	55.53	27.05	68.91
CB	53.13	28.52	67.78
CA	53.11	28.76	64.74
C	50.19	28.81	63.81
O	48.38	27.92	65.25
N	54.58	31.15	63.85

126 TYR

OH	45.77	37.38	70.85
CZ	46.14	32.48	65.40
CE2	46.21	33.54	66.05
CZ	45.85	36.33	68.30
CE1	45.69	37.99	66.19
CD1	45.82	36.84	63.49
CG	45.96	34.09	63.20
CB	46.16	32.95	60.35
CA	47.36	30.02	60.29
C	46.76	28.96	57.44
O	44.61	29.18	56.24
N	49.79	29.85	61.42

127 GLY

CA	48.86	26.73	53.08
C	48.89	28.50	51.19
O	49.37	27.55	48.95
N	48.91	27.82	56.41

128 ASP

OD2	44.16	28.15	47.50
OD1	42.33	31.91	47.76
CG	44.34	30.55	47.46
CB	47.07	31.63	46.89
CA	48.31	32.97	49.39
C	46.75	35.52	50.32
O	45.55	35.78	52.17
N	48.37	31.06	51.61

129 PHE

CD2	45.49	45.78	49.39
CE2	45.03	48.66	49.78
CZ	43.83	50.03	47.79
CE1	43.37	49.81	45.41
CD1	44.23	48.12	45.03
CG	45.56	44.76	46.96
CB	46.37	41.92	46.45
CA	45.36	39.92	48.49
C	42.34	39.51	48.22
O	40.77	40.51	49.86
N	46.78	37.37	48.13

130 GLY

CA	38.73	37.52	45.67
C	37.52	36.08	48.08
O	35.10	36.63	48.58
N	41.58	38.05	46.16

131 LYS

NZ	32.50	24.27	52.41
CE	34.01	26.55	51.29
CD	36.56	27.04	52.88
CG	38.78	28.17	51.12
CB	39.47	30.68	52.34
CA	38.32	33.21	51.91
C	38.23	34.24	54.77
O	36.85	33.19	56.51
N	39.22	34.69	49.54

132 ALA

CB	42.72	38.39	58.54
CA	39.80	37.67	57.88
C	38.24	40.30	57.94
O	37.92	41.54	60.06
N	39.70	38.44	55.22

133 VAL

CG2	34.27	47.15	52.15
CG1	34.87	42.41	50.43
CB	35.75	44.55	52.49
CA	35.78	43.62	55.39
C	32.97	43.05	56.45
O	31.08	44.58	55.94
N	37.30	41.12	55.61

134 GLN

NE	32.64	32.42	57.73
CE	32.03	32.64	53.29
CD	31.75	33.58	55.63
CG	30.12	30.16	55.97
CB	29.83	37.09	56.74
CA	30.13	40.07	59.03
C	29.57	40.79	61.95
O	27.25	40.88	62.82
N	32.72	40.86	57.90

135 GLN

NE	33.83	33.40	71.60
CE	34.04	32.96	67.85
CD	33.39	35.06	69.11
CG	31.84	37.24	67.60
CB	32.94	39.95	67.87
CA	31.45	42.31	66.20
C	32.41	44.76	67.19
O	33.36	46.40	65.58
N	31.71	41.30	63.41

136 PRO

CD	31.55	43.24	71.82
CG	32.51	44.52	74.42
CB	32.41	47.50	73.86
CA	32.99	47.80	70.94
C	35.94	48.30	70.31
O	36.64	50.28	69.00
N	32.16	45.23	69.78

137 ASP

OD2	41.05	41.73	74.61
OD1	44.77	42.11	72.75
CG	42.69	43.11	73.54
CB	42.04	46.05	73.41
CA	40.52	46.76	70.82
C	42.29	45.35	66.76
O	44.38	44.20	69.45
N	37.62	46.50	71.27

138 GLY

CA	42.93	44.16	64.10
C	45.14	45.49	63.03
O	47.37	45.09	62.43
N	41.43	45.45	66.27

134 LEU

CD2	43.44	53.15	54.88
CG1	45.47	48.07	56.25
CG	45.74	51.64	56.72
CB	45.11	52.26	54.62
CA	46.42	53.52	61.80
C	47.40	52.65	63.42
O	47.16	53.25	65.70
N	44.48	48.54	62.81

140 ALA

CB	54.27	54.65	64.48
CA	51.58	53.62	63.54
C	52.26	57.95	61.40
O	53.56	57.36	59.42
N	49.97	53.75	62.21

141 VAL

CG2	47.36	61.61	56.04
CG1	49.47	66.40	57.55
CB	49.21	63.64	54.27
CA	51.85	62.61	60.14
C	53.65	64.09	61.50
O	52.96	65.76	63.63
N	51.36	60.38	61.99

142 LEU

CD2	62.13	70.33	63.31
CD1	65.44	66.85	62.14
CG	62.66	67.37	63.27
CB	60.61	65.92	61.59
CA	57.81	67.19	61.35
C	57.81	69.59	59.46
O	58.67	69.39	57.14
N	55.91	65.23	60.25

143 GLY

CA	56.73	74.29	58.87
C	58.70	76.45	59.78
O	59.67	76.41	62.37
N	56.85	71.83	60.48

144 ILE

CB	67.21	75.92	57.99
CG1	64.48	76.91	56.93
CB	63.95	74.93	57.44
CG2	65.76	81.29	55.51
CA	61.14	80.48	58.60
C	60.13	83.11	57.41
O	59.93	83.41	54.96
N	59.26	78.30	57.29

145 PHE

CD2	54.24	87.60	64.18
CE2	51.56	86.83	64.38
CZ	49.54	87.03	63.12
CE1	50.35	87.83	60.58
CD1	52.72	88.51	59.84
CG	54.75	88.46	61.63
CB	57.53	89.16	60.71
CA	58.30	87.61	58.24
C	60.88	88.90	56.86
O	62.47	90.41	58.09
N	59.51	84.99	59.17

146 LEU

CD2	60.76	84.61	44.70
CD1	62.57	87.21	45.88
CG	61.51	87.19	48.70
CB	61.61	86.61	50.50
CA	62.65	93.60	50.84
C	61.41	93.14	51.52
O	59.84	93.09	50.25
N	60.53	89.38	54.29

147 LYS

NZ	59.57	105.09	44.74
CE	59.45	102.08	56.07
CD	62.96	101.11	54.41
CG	60.77	99.72	51.26
CB	61.75	97.95	50.98
CA	63.42	97.47	49.45
C	63.42	98.60	50.31
O	66.25	99.80	50.31
N	62.74	95.37	51.97

148 VAL

CG2	66.42	101.44	41.04
CG1	65.31	100.32	44.15
CB	63.60	100.44	44.14
CA	65.04	102.10	45.52
C	67.14	103.05	47.29
O	66.42	104.41	49.11
N	63.14	100.59	47.19

149 GLY

CA	72.15	103.07	48.03
C	74.04	103.57	46.53
O	74.04	105.27	44.52
N	69.76	102.32	46.50

150 SER

OG	81.51	103.15	48.73
CB	81.58	102.65	47.36
CA	77.23	102.35	45.45
C	77.33	99.76	45.91
O	77.30	99.11	43.74
N	76.73	102.07	46.92

151 ALA

CB	84.35	97.27	39.40
CA	81.83	97.36	40.04
C	81.67	94.52	42.35
O	82.33	94.56	44.74
N	81.50	99.45	42.31

152 LYS

NZ	77.03	91.70	48.17
CE	74.25	94.42	49.39
CD	74.43	97.47	47.63
CG	77.35	97.11	45.52
CB	77.84	99.29	43.34
CA	80.33	99.34	42.52
C	81.53	97.93	40.11
O	80.15	97.44	38.34
N	80.74	92.37	41.10

153 PHE

CD	88.67	88.50	42.33
CG	88.58	88.40	41.65
CB	88.51	88.05	38.63
CA	85.62	88.59	37.88
C	84.44	83.76	37.71
O	83.43	82.31	35.39
N	84.28	87.99	40.09

154 GLY

CA	81.37	74.72	39.42
C	79.11	74.53	39.02
O	78.71	77.41	39.11
N	84.14	82.46	39.89

155 LEU

CD2	72.73	88.20	40.72
CD1	72.20	91.58	42.66
CG	71.77	91.57	39.85
CB	74.41	88.21	38.88
CA	76.10	91.44	37.25
C	78.63	91.15	37.22
O	79.33	91.04	32.84
N	79.09	81.82	38.16

156 GLN

NE2	78.75	93.39	25.92
CD	79.84	94.42	24.26
CG	78.47	94.27	27.52
CB	77.47	94.44	28.03
CA	77.73	94.80	30.37
C	77.44	92.27	28.65
O	76.61	91.84	26.11
N	77.78	84.19	33.24

157 THR

CG2	80.63	84.47	17.03
CB	82.14	84.24	17.44
OG1	83.77	84.94	21.15
CA	77.75	84.18	19.19
C	76.34	84.10	17.57
O	73.10	84.10	16.38
N	78.40	84.27	15.18

158 VAL

CG2	71.44	81.34	16.36
CG1	71.43	81.43	15.45
CB	73.64	80.72	14.80
CA	75.12	80.40	13.18
C	76.12	80.40	12.21
O	77.14	78.87	13.44

159 VAL

CG2	88.44	84.66	30.89
CG1	84.46	86.42	26.38
CB	84.89	84.56	27.95
CA	84.06	81.81	26.76
C	84.49	81.76	23.71
O	87.64	80.78	22.26
N	76.60	79.70	27.70

160 ASP

CG	78.94	85.37	19.37
OD1	75.36	87.53	19.40
CB	76.53	85.34	19.35
CB	75.21	82.69	19.42
CA	72.17	82.94	19.91
C	70.74	80.78	18.28
O	69.22	81.37	16.42
N	71.70	82.20	23.10

161 VAL

CG2	71.16	71.13	17.44
CG1	68.12	72.81	21.15
CB	70.43	73.48	19.19
CA	70.04	76.33	17.57
C	67.05	76.33	16.38
O	65.97	75.02	15.18
N	71.00	78.10	18.70

162 LEU

CD2	59.15	78.31	23.04
CD1	63.43	82.97	24.10
CG	61.32	80.40	22.80
CB	61.98	80.92	15.89
CA	62.69	78.53	18.16
C	62.16	76.08	15.24
O	59.80	79.42	14.56
N	65.70	77.89	18.39

163 ASP

CG	64.63	82.54	7.72
OD1	65.81	84.20	8.02
OD2	67.29	82.28	8.22
CB	66.47	79.56	9.21
CA	63.77	79.68	10.70
C	61.80	77.62	9.59
O	59.43	78.21	9.18
N	64.30	79.20	13.50

164 SER

OG	65.23	70.53	8.76
CB	62.39	70.37	9.58
CA	61.13	73.07	8.34
C	58.23	72.66	8.91
O	56.43	72.34	7.23
N	62.80	75.20	8.90

165 ILE

CG2	55.42	66.23	15.82
CG1	55.22	68.44	15.57
CB	52.76	72.62	16.73
CA	53.12	72.29	12.62
C	50.82	74.34	12.91
O	49.77	74.07	14.51
N	57.82	72.67	11.52

166 LYS

NZ	54.23	63.26	9.25
CE	56.28	65.61	9.57
CD	54.91	67.83	9.98
CG	53.96	69.89	8.67
CB	53.96	69.09	8.76
CA	56.64	70.69	11.46
C	57.63	73.25	12.23
O	58.84	73.44	14.74
N	54.18	76.52	11.37

167 THR

OG1	49.24	77.13	7.03
CB	51.04	72.91	7.45
CA	52.48	73.11	8.97
C	54.56	75.08	10.64
O	54.12	72.33	12.93
N	55.14	76.40	13.44

168 LYS

NZ	33.70	61.46	16.32
CE	33.15	63.50	17.77
CD	35.91	65.91	16.18
CG	37.17	66.02	16.36
CB	38.14	65.84	15.45
CA	41.37	67.24	14.80
C	41.60	66.43	13.18
O	39.76	65.13	12.21
N	41.64	71.63	13.44

169 GLY

CA	44.42	63.26	11.80
C	47.99	62.52	11.34
O	48.74	60.35	12.21
N	44.14	65.72	13.01

170 LYS

N	52.37	72.17	0.70
CE	51.03	71.00	0.08
CD	52.68	68.96	9.20
CG	51.81	68.39	7.36
CB	33.07	80.32	8.65
CA	52.28	63.85	9.77
C	53.97	63.35	12.18
O	53.58	64.07	14.41
N	49.40	64.32	10.15

171 SER

OG	54.17	56.82	14.05
CB	56.89	57.44	14.04
CA	57.60	60.28	13.88
C	60.52	60.35	12.99
O	61.16	59.02	10.02
N	55.83	61.23	11.71

172 ALA

CB	60.09	63.96	14.85
CA	65.14	61.08	14.34
C	66.74	59.30	16.24
O	66.94	59.00	18.60
N	62.26	60.91	14.91

173 ASP

OD2	70.05	51.50	13.03
OD1	71.01	54.57	10.36
CG	71.08	53.62	12.61
CB	72.53	54.65	14.95
CA	70.58	56.36	16.80
C	72.07	58.09	16.85
O	73.66	59.56	18.12
N	68.87	58.15	15.19

174 PHE

CD2	71.84	62.66	25.49
CE2	72.44	65.40	30.69
CZ	72.73	67.63	29.20
CE1	72.32	67.52	26.53
CD1	71.63	65.02	25.32
CG	71.47	62.74	26.77
CB	70.71	60.19	25.39
CA	72.84	59.05	23.55
C	74.78	57.30	25.13
O	74.20	56.50	27.40
N	71.54	57.55	21.38

175 THR

CG	81.73	50.82	24.37
OG	77.56	51.65	22.19
CB	80.33	52.96	23.15
CA	79.10	55.07	25.21
C	81.68	56.16	26.42
O	83.66	55.64	25.00
N	77.06	56.75	23.91

176 ASN

ND	88.09	53.61	28.74
OD	87.25	55.30	24.65
CG	87.17	55.62	27.06
CB	86.26	58.14	28.43
CA	84.00	57.58	30.43
C	84.23	60.05	32.24
O	85.46	62.06	31.67
N	81.64	56.53	29.04

177 PHE

CD2	77.50	65.98	39.76
CE2	76.37	68.61	39.54
CZ	77.06	70.28	37.54
CE1	78.76	69.42	35.61
CD1	79.82	66.78	35.76
CG	79.25	65.13	37.82
CB	80.41	62.36	37.84
CA	83.12	62.10	36.49
C	84.76	62.91	38.95
O	87.19	63.19	38.79
N	83.09	59.87	34.59

178 ASP

OD2	88.91	63.79	47.18
OD1	91.98	64.32	44.40
CG	89.61	64.36	44.96
CB	87.42	65.24	43.09
CA	84.76	63.80	43.72
C	82.49	65.51	44.86
O	82.73	67.97	45.08
N	83.42	63.04	41.22

179 PRO

CD	80.08	61.29	45.80
CG	77.84	60.84	47.59
CB	76.00	63.40	47.19
CA	77.98	65.54	48.65
C	78.80	67.14	47.16
O	77.15	67.64	50.94
N	80.31	64.19	45.55

180 ARG

NH2	90.32	59.28	56.73
NH1	90.56	63.68	55.45
CZ	89.15	61.41	55.42
NE	88.83	61.24	54.05
CD	85.75	63.49	52.43
CG	83.30	65.92	54.11
CB	83.66	66.42	52.31
CA	82.39	64.37	51.50
C	82.69	72.41	51.48
O	82.03	73.81	53.42
N	81.33	67.87	49.25

181 GLY

CA	84.10	78.29	48.92
C	81.20	77.27	48.90
O	80.82	74.54	48.08
N	83.69	73.40	49.24

182 LEU

CD2	72.44	78.23	48.74
CD1	70.34	73.94	47.57
CG	72.61	75.22	46.37
CB	74.33	74.18	48.21
CA	76.56	76.22	49.86
C	75.64	78.23	52.78
O	73.24	78.47	53.36
N	79.41	75.53	49.79

183 LEU

CD2	76.24	69.07	61.42
CD1	76.75	68.92	57.08
CG	75.20	71.04	59.16
CB	76.08	73.16	58.75
CA	76.95	75.93	57.45
C	77.91	78.21	59.25
O	80.15	79.17	58.99
N	77.56	75.96	54.58

184 PRO

CD	73.71	77.87	61.81
CG	73.04	78.79	64.57
CB	74.41	81.50	64.62
CA	76.85	81.21	62.92
C	79.37	80.49	64.49
O	80.57	78.30	64.07
N	76.18	79.03	61.07

185 GLU

OE2	89.52	85.09	73.10
OE1	85.22	86.30	73.25
CD	87.22	85.31	72.25
CG	87.11	83.96	69.40
CB	84.42	84.29	67.92
CA	82.53	81.82	67.93
C	81.92	81.21	70.87
O	83.74	83.66	72.46
N	80.13	82.26	66.29

186 SER

OG	75.27	82.80	77.44
CB	77.27	83.24	75.51
CA	78.51	80.77	74.29
C	76.54	78.47	73.84
O	74.81	76.59	72.07
N	79.36	81.30	71.53

187 LEU

CD2	80.90	71.00	73.82
CD1	78.65	69.96	71.52
CG	78.11	69.30	74.16
CB	76.57	71.32	76.20
CA	75.09	74.04	75.23
C	72.52	74.39	76.84
O	70.68	72.73	76.06
N	76.84	76.39	75.45

188 ASP

OD2	73.35	78.39	83.93
OD1	73.13	82.56	83.88
CG	72.38	80.40	83.35
CB	70.17	80.01	81.04
CA	70.07	77.09	80.04
C	67.48	76.59	78.48
O	67.41	76.89	76.02
N	65.41	75.84	79.93
	72.42	74.53	78.38

189 TYR

OH	64.87	64.41	85.20
CD2	65.93	69.31	79.86
CE2	66.32	67.26	81.77
CZ	64.36	66.36	83.32
CE1	61.84	67.23	82.96
CD1	61.31	69.32	81.01
CG	63.46	70.26	79.46
CB	62.89	72.49	77.40
CA	62.81	73.29	78.68
C	60.06	73.29	80.01
O	59.34	75.16	82.47
N	65.41	75.84	79.93

190 TRY

CD2	55.03	83.08	76.36
CE2	55.21	83.73	75.09
CZ	55.85	86.42	75.16
CH	56.39	88.32	77.11
CZ1	56.17	87.73	79.77
CE1	55.50	85.09	80.32
NE	55.17	84.07	82.63
CD1	54.60	81.24	82.28
CG	54.43	80.68	79.67
CB	53.81	77.93	78.27
CA	55.25	75.48	79.38
C	54.05	72.81	78.50
O	54.72	71.77	76.34
N	58.01	75.44	78.36

191 THR

CG	51.90	66.61	83.86
OG	52.10	64.60	79.49
CB	52.69	66.96	80.97
CA	51.05	69.16	79.57
C	48.13	69.21	80.48
O	47.48	70.28	82.61
N	52.33	71.73	80.17

192 TYR

OH	38.14	74.34	69.50
CD2	42.97	72.66	74.86
CE2	41.93	73.51	72.34
CZ	39.16	73.27	71.95
CE1	37.47	72.42	73.87
CD1	38.56	71.55	76.44
CG	41.33	71.78	76.80
CB	42.45	70.84	79.57
CA	43.54	67.96	79.47
C	41.92	66.30	77.49
O	42.64	66.07	75.13
N	46.41	68.05	78.84

193 PRO

CD	38.52	65.78	81.05
CG	35.64	64.80	80.76
CB	35.79	62.67	78.66
CA	38.00	63.52	76.77
C	37.05	65.16	74.37
O	36.15	67.45	74.66
N	39.74	65.18	78.47

194 GLY

CA	36.43	65.31	69.56
C	34.89	63.69	67.47
O	33.96	61.46	68.01
N	37.26	63.96	72.03

195 SER

OG	28.51	62.44	62.04
CB	30.32	64.50	62.79
CA	33.18	63.52	62.93
C	34.58	64.04	60.26
O	36.42	65.69	60.09
N	34.66	64.90	65.10

196 LEU

CD2	40.17	58.56	51.37
CD1	39.13	60.16	53.01
CG	38.43	59.53	51.38
CB	34.27	60.45	53.83
CA	34.86	62.94	55.54
C	33.26	65.50	55.02
O	31.13	65.93	56.21
N	33.57	62.77	58.16

197 THR

CG	35.75	72.88	55.36
OG	37.41	70.62	51.45
CB	35.13	71.08	52.59
CA	32.98	69.08	52.58
C	31.48	69.72	49.91
O	30.50	71.82	49.02
N	34.46	66.10	51.72

198 THR

CO	31.17	70.01	92.58
CO	34.41	63.30	95.27
CB	31.44	64.01	93.90
CA	29.45	67.15	96.08
C	29.10	64.21	96.04
O	30.24	62.54	97.97
N	29.52	68.19	91.36

199 PRO

CD	26.49	66.18	93.07
CG	25.20	64.41	91.35
CB	25.32	61.19	92.39
CA	26.63	61.71	94.43
C	26.15	59.31	94.10
O	30.49	59.53	93.77
N	27.73	63.41	90.74

200 PRO

CD	29.31	54.57	94.73
CG	26.17	52.17	94.[illegible]
CB	25.70	53.50	97.33
CA	25.43	56.57	96.67
C	25.39	59.12	97.67
O	27.39	59.33	99.39
N	26.98	63.90	93.52

201 LEU

CD2	25.00	60.34	91.59
CD1	23.91	63.00	91.12
CG	23.93	63.54	91.24
CB	24.82	61.77	93.53
CA	23.28	59.14	94.00
C	24.04	57.67	96.56
O	23.85	56.75	98.78
N	29.97	67.18	94.72

202 LEU

CD2	28.26	50.34	92.73
CD1	23.87	49.83	94.94
CG	26.86	49.56	93.29
CB	27.76	51.35	97.57
CA	25.85	53.59	98.50
C	27.24	55.37	100.58
O	27.33	57.72	100.05
N	25.02	55.23	96.21

203 GLU

OE2	21.93	50.88	103.71
OE1	21.53	50.35	100.03
CD	22.32	50.17	100.65
CG	24.06	50.43	100.70
CB	27.06	50.32	97.41
CA	29.03	53.72	97.08
C	31.38	54.12	100.20
O	31.62	53.74	101.49
N	27.72	54.23	102.91

204 CYH

SG	35.17	44.37	99.11
CB	35.66	49.15	98.53
CA	35.43	51.61	97.22
C	38.37	52.46	97.27
O	40.17	50.83	98.74
N	33.08	53.20	94.41

205 VAL

CG2	41.84	48.69	89.40
CG1	39.49	50.35	88.23
CB	42.06	51.90	88.68
CA	41.57	54.03	90.03
C	42.53	54.74	92.04
O	40.92	54.82	94.33
N	38.83	54.97	89.92

206 THR

OG1	44.84	58.15	87.[illegible]
CG2	46.03	55.37	87.[illegible]
CB	47.76	56.[illegible]	87.[illegible]
CA	46.30	58.40	87.[illegible]
C	47.92	61.[illegible]	86.07
O	49.[illegible]	61.[illegible]	87.[illegible]
N	45.13	57.50	88.96

207 TRY

CE2	44.75	64.[illegible]	84.[illegible]
CE1	46.[illegible]	62.[illegible]	84.[illegible]
CZ2	46.12	64.93	81.37
CH2	45.57	64.55	80.03
CZ1	47.[illegible]	62.60	83.[illegible]
CE1	46.[illegible]	61.[illegible]	84.[illegible]
NE	46.[illegible]	67.31	86.35
CD1	46.[illegible]	61.[illegible]	86.56
CD	44.[illegible]	63.[illegible]	80.11
CE	47.[illegible]	66.28	78.13
CA	48.22	[illegible]	[illegible]
C	51.[illegible]	[illegible]	71.52
O	51.36	66.11	[illegible]
N	47.67	63.56	71.23

208 ILE

CD1	57.[illegible]	61.60	73.05
CG2	57.[illegible]	[illegible]	75.[illegible]
CG1	55.57	63.[illegible]	73.74
CB	55.27	63.[illegible]	73.[illegible]
CA	56.22	59.15	71.38
C	56.77	57.07	69.[illegible]
O	57.[illegible]	57.[illegible]	68.98
N	53.69	57.02	70.10

209 VAL

CG2	52.44	53.79	63.[illegible]
CG1	53.[illegible]	55.[illegible]	66.28
CB	54.45	54.[illegible]	64.[illegible]
CA	54.[illegible]	53.49	64.50
C	54.[illegible]	51.[illegible]	63.[illegible]
O	55.87	51.[illegible]	63.26
N	52.58	53.23	63.[illegible]

210 LEU

CD2	48.43	54.54	64.30
CD1	48.43	52.[illegible]	64.44
CG	48.44	53.[illegible]	65.[illegible]
CB	48.[illegible]	54.[illegible]	67.[illegible]
CA	48.[illegible]	54.[illegible]	68.[illegible]
C	48.[illegible]	54.72	70.[illegible]
O	48.[illegible]	53.[illegible]	70.[illegible]
N	48.[illegible]	52.[illegible]	68.[illegible]

211 LYS

NZ	45.91	42.44	83.55
CE	44.77	43.54	81.[illegible]
CD	46.[illegible]	44.[illegible]	84.30
CG	46.63	[illegible]	77.27
CB	46.07	47.23	76.53
CA	46.83	[illegible]	73.57
C	47.[illegible]	51.20	73.[illegible]
O	45.63	52.[illegible]	71.79
N	43.92	43.60	73.16

212 GLU

OE2	74.44	50.35	[illegible]
OE1	74.54	[illegible]	69.78
CD	74.13	48.76	70.63
CG	73.[illegible]	[illegible]	[illegible]
CB	73.[illegible]	[illegible]	71.[illegible]
CA	70.13	48.27	70.04
C	70.27	[illegible]	[illegible]
O	71.54	50.16	69.89
N	68.13	46.89	71.41

213 PRO

CD	66.62	[illegible]	[illegible]
CG	65.98	[illegible]	[illegible]
CB	66.47	48.50	[illegible]
CA	67.66	[illegible]	[illegible]
C	70.[illegible]	48.40	61.39
O	71.[illegible]	[illegible]	62.[illegible]
N	68.14	49.06	65.82

214 ILE

CD1	71.65	[illegible]	[illegible]
CG2	73.[illegible]	[illegible]	[illegible]
CG1	73.25	50.[illegible]	[illegible]
CB	74.35	49.10	57.80
CA	72.66	47.45	57.38
C	71.[illegible]	[illegible]	[illegible]
O	68.97	47.08	[illegible]
N	70.12	48.30	58.80

215 SER

OG	76.[illegible]	[illegible]	[illegible]
CB	77.[illegible]	45.07	53.57
CA	72.58	[illegible]	52.65
C	[illegible]	[illegible]	[illegible]
O	75.63	[illegible]	49.07
N	71.27	47.22	54.[illegible]

216 VAL

CG2	69.[illegible]	45.[illegible]	[illegible]
CG1	70.[illegible]	[illegible]	45.41
CB	71.[illegible]	[illegible]	[illegible]
CA	72.[illegible]	[illegible]	44.40
C	72.[illegible]	[illegible]	42.37
O	[illegible]	133.39	[illegible]
N	72.09	[illegible]	47.75

217 SER

OG	70.23	[illegible]	38.42
CB	70.32	133.53	36.59
CA	74.15	[illegible]	[illegible]
C	71.[illegible]	[illegible]	37.33
O	71.21	[illegible]	35.23
N	74.24	98.74	40.98

218 SER

OG	87.[illegible]	[illegible]	[illegible]
CB	87.57	102.23	38.44
CA	90.[illegible]	[illegible]	33.18
C	70.00	103.26	38.31
O	[illegible]	[illegible]	[illegible]
N	[illegible]	[illegible]	[illegible]

219 GLU

OE2	73.05	[illegible]	24.45
OE1	70.41	[illegible]	27.05
CD	72.77	104.32	28.31
CG	74.[illegible]	[illegible]	29.65
CB	73.32	99.29	[illegible]
CA	[illegible]	[illegible]	31.20
C	69.74	97.00	32.18
O	72.33	96.33	27.87
N	69.97	101.46	33.62

220 GLN

OE1	70.47	[illegible]	35.36
NE2	72.54	[illegible]	37.55
CD	73.43	[illegible]	35.42
CG	74.42	[illegible]	36.74
CB	69.53	90.77	32.83
CA	[illegible]	92.61	31.19
C	66.[illegible]	93.14	[illegible]
O	63.17	[illegible]	30.71
N	68.58	95.38	31.91

221 VAL

CG2	61.38	87.33	34.03
CG1	63.42	100.74	35.38
CB	61.39	98.18	36.31
CA	[illegible]	97.50	34.34
C	60.43	96.58	31.70
O	58.65	95.10	[illegible]
N	64.54	95.16	33.55

222 LEU

CD2	60.04	104.76	30.29
CD1	62.23	106.74	28.28
CG	62.27	[illegible]	29.22
CB	61.84	101.56	26.71
CA	60.11	[illegible]	27.98
C	59.59	97.83	25.72
O	[illegible]	97.75	[illegible]
N	61.27	98.77	30.53

223 LYS

NZ	[illegible]	[illegible]	18.21
CE	[illegible]	94.82	18.46
CD	[illegible]	[illegible]	22.33
CG	[illegible]	[illegible]	23.37
CB	[illegible]	92.27	22.09
CA	[illegible]	91.95	23.13
C	[illegible]	[illegible]	23.80
O	58.83	[illegible]	24.83
N	61.53	95.45	25.26

224 PHE

CD1	[illegible]	86.30	11.93
CD2	[illegible]	[illegible]	15.23
CE1	62.98	84.33	14.32
CE2	63.36	82.[illegible]	[illegible]
CZ	[illegible]	83.74	13.30
CG	[illegible]	87.29	11.96
CB	58.43	89.12	[illegible]
CA	58.99	[illegible]	16.93
C	[illegible]	89.79	19.75
O	[illegible]	[illegible]	[illegible]
N	59.15	93.78	16.06

225 ARG

NE	[illegible]	[illegible]	[illegible]
CZ	[illegible]	102.02	33.61
NH2	[illegible]	[illegible]	34.66
NH1	49.82	99.06	33.19
CD	[illegible]	[illegible]	33.25
CG	[illegible]	[illegible]	29.17
CB	50.25	[illegible]	[illegible]
CA	[illegible]	[illegible]	25.43
C	[illegible]	93.17	24.49
O	[illegible]	94.46	[illegible]
N	53.34	[illegible]	26.51

226 LYS

NZ	[illegible]	[illegible]	[illegible]
CE	57.97	[illegible]	[illegible]
CD	57.41	95.39	[illegible]
CG	56.66	[illegible]	15.40
CB	[illegible]	[illegible]	[illegible]
CA	[illegible]	[illegible]	[illegible]
C	[illegible]	[illegible]	[illegible]
O	53.33	[illegible]	[illegible]
N	[illegible]	93.61	15.06
	[illegible]	91.77	20.95

227 LEU

CD2	[illegible]	81.93	24.04
CD1	[illegible]	[illegible]	[illegible]
CG	[illegible]	[illegible]	[illegible]
CB	48.31	[illegible]	[illegible]
CA	48.76	94.54	[illegible]
C	[illegible]	[illegible]	[illegible]
O	44.57	90.92	[illegible]
N	49.95	87.24	19.28

228 ASN

OD1	[illegible]	90.35	10.35
ND2	[illegible]	[illegible]	11.66
CG	[illegible]	89.20	11.71
CB	[illegible]	87.27	[illegible]
CA	41.99	86.91	16.07
C	[illegible]	[illegible]	17.30
O	[illegible]	78.10	[illegible]
N	44.77	82.54	16.99

	229 PHE				234 VAL				240 PRO		
CL	35.16	74.71	24.71	CE	42.48	73.10	24.03	CD	1.71	73.37	88.88
CE1	35.93	80.98	27.15	CG	43.25	74.11	25.00	CG	2.63	75.15	90.67
CZ	34.72	83.49	27.75	C	41.71	82.44	23.00	C	1.41	81.77	73.12
CE1	33.14	84.79	28.23	O	41.03	83.52	23.15	CA	0.11	82.13	73.17
CD1	33.25	83.58	25.29	CA	39.32	81.18	23.70	C	0.67	82.67	74.38
CG	34.18	81.49	24.74	O	40.48	87.41	25.74	O	23.76	73.13	73.84
CB	33.83	79.96	24.40	CB	42.40	81.62	25.88	N	21.95	81.86	70.36
CA	36.25	77.61	23.17	N	38.68	87.40	21.11				
C	35.43	76.71	16.40						241 ASP		
O	34.50	74.34	17.03		235 VAL			CG	4.49	82.71	86.07
N	37.81	80.66	17.60	CG1	37.75	94.37	26.81	OD1	4.41	77.74	81.52
				CG2	42.83	94.31	26.47	OD	24.15	83.38	81.34
	230 ASP			CB	40.30	93.44	28.19	CB	25.36	75.33	81.57
OD1	27.93	77.01	23.23	CA	40.01	91.47	24.11	CA	23.99	78.73	74.36
OD2	30.14	80.73	21.31	C	38.20	91.16	31.72	C	22.16	78.49	80.27
CG	29.50	78.49	22.23	O	36.34	91.71	31.60	O	23.44	74.15	80.81
CB	30.87	79.40	23.53	N	38.43	89.34	27.28	N	22.59	80.55	77.48
CA	33.46	79.36	23.43								
C	33.61	74.36	23.33		241 ASP				250 LYS		
O	32.71	73.30	22.57	OD2	38.31	97.31	30.04	NZ	13.84	74.50	86.66
N	34.43	77.22	24.45	OD1	35.21	96.44	33.19	CE	12.35	76.94	86.71
				CG	36.63	97.07	35.07	CD	11.74	76.78	89.53
	231 GLY			CB	33.13	94.43	36.13	CG	10.93	78.07	87.44
CA	30.33	73.61	26.16	CA	37.38	93.66	36.33	CB	11.91	79.29	86.46
C	30.55	74.60	24.50	C	38.01	94.84	37.32	CA	11.17	79.24	84.16
O	30.94	75.71	24.93	O	33.57	94.93	37.41	C	11.15	78.05	85.15
N	34.78	73.82	26.38	N	38.99	92.40	33.37	O	13.03	75.39	87.06
								N	20.00	77.23	82.50
	232 GLU				242 ASN						
OE2	41.57	73.75	31.74	ND2	41.83	94.72	37.32		251 ASN		
OE1	43.43	75.11	30.43	OD1	42.33	91.44	34.71	ND2	14.93	69.04	93.28
CD	42.38	73.48	31.51	CG	40.83	92.33	36.07	OD1	19.53	64.64	92.06
CG	40.78	72.05	29.17	CB	37.83	92.77	33.42	CG	20.23	66.80	91.74
CB	37.79	72.23	29.50	CA	36.53	94.47	37.79	CB	21.49	67.21	89.33
CA	36.35	74.27	31.84	C	36.45	93.71	40.06	CA	22.43	70.12	89.08
C	33.39	74.36	31.35	O	36.10	91.77	41.14	C	23.33	69.36	87.05
O	32.71	75.27	31.53	N	37.61	87.54	37.15	O	26.04	67.44	89.79
N	36.40	72.99	31.11					N	20.70	71.53	86.78
					243 TRP						
	233 GLY			CD2	34.24	91.39	47.84		252 ARG		
CA	28.73	73.75	34.35	CE2	32.81	93.11	50.03	NH2	30.98	69.30	75.96
C	27.75	73.43	33.24	CZ2	30.13	92.80	50.98	NH1	28.15	73.13	76.32
O	25.71	74.66	33.24	CH2	30.27	94.36	49.40	CZ	29.35	70.99	77.41
N	31.68	73.59	33.56	CZ3	31.74	95.40	47.86	NE	28.72	70.33	80.00
				CE3	33.73	94.69	46.89	CD	30.56	68.36	81.45
	234 GLU			NE	35.36	94.21	49.00	CG	29.70	69.16	84.35
OE2	23.91	73.48	37.74	CD1	37.03	94.80	49.58	CB	31.31	70.15	85.09
OE1	23.88	75.32	36.41	CG	38.43	93.30	48.42	CA	32.02	70.82	86.80
CD	23.05	74.38	36.35	CD	37.84	92.13	46.83	C	32.43	72.41	89.79
CG	25.45	77.25	37.13	CA	36.25	94.43	45.38	O	32.73	74.81	89.23
CB	26.28	76.03	36.77	C	33.07	94.17	46.85	N	27.13	71.28	88.53
CA	26.43	74.50	34.82	O	31.54	93.81	45.66				
C	30.48	81.74	34.14	N	35.99	89.51	42.53		253 GLN		
O	32.42	81.58	33.33					NE2	39.41	73.77	100.35
N	29.17	77.66	33.86		244 ARG			OE1	40.13	71.39	99.88
				NH2	26.54	72.69	56.75	CD	38.86	73.50	99.26
	235 PRO			NH1	22.96	72.83	54.71	CG	38.03	73.10	97.19
CD	27.57	84.31	31.07	CZ	24.75	73.44	55.11	CB	37.04	71.28	97.10
CG	27.71	87.29	30.42	NE	25.43	76.10	53.85	CA	36.58	72.33	92.32
CB	30.06	87.71	30.44	CD	28.03	77.57	54.07	C	39.13	71.95	90.67
CA	31.69	85.06	32.43	CG	28.86	74.96	51.72	O	39.55	69.83	89.45
C	34.14	87.28	33.24	CB	31.41	75.07	51.96	N	34.17	71.04	91.24
O	33.27	88.25	35.34	CA	31.91	73.65	53.20				
N	30.04	83.78	32.53	C	31.42	69.63	53.81		254 ILE		
				O	33.27	69.16	55.46	CD	43.56	77.01	82.49
	236 GLU			N	33.84	84.25	49.32	CD1	45.25	75.35	84.58
OE2	28.03	92.50	41.11					CB	43.43	70.11	86.98
OE1	23.96	91.34	43.37		245 PRO			CG2	44.29	78.55	88.62
CD	25.84	91.53	42.00	CD	28.34	86.63	52.26	CA	43.31	73.40	84.15
CG	25.98	90.30	41.10	CG	29.71	87.91	54.34	C	45.93	74.25	90.69
CB	26.37	91.57	39.67	CB	30.69	89.72	53.19	O	47.97	73.44	92.37
CA	31.23	91.04	39.11	CA	28.93	89.39	50.57	N	40.79	74.00	90.66
C	32.76	89.73	41.35	C	29.03	88.80	49.20				
O	34.31	90.91	42.03	O	29.21	89.41	48.14		255 LYS		
N	31.09	89.68	39.33	N	29.28	86.75	54.15	NZ	48.55	83.06	89.21
								CE	46.84	84.07	91.43
	237 GLU				246 ALA			CD	48.45	86.59	91.09
OE2	35.83	85.72	7.16	CB	31.27	93.09	55.99	CG	49.25	87.05	93.10
OE1	34.72	81.49	8.47	CA	30.40	89.62	53.71	CB	51.59	86.73	92.04
CD	35.49	83.96	7.93	C	27.82	89.19	54.06	CA	50.71	83.40	90.83
CG	36.18	85.32	8.33	O	25.93	90.47	53.61	C	52.64	84.27	88.82
CB	33.71	83.44	12.49	N	30.14	87.72	51.19	O	52.42	84.18	86.38
CA	32.75	80.27	13.73					N	48.07	83.21	89.55
C	35.22	87.40	15.14		247 LEU						
O	37.43	88.31	14.72	CE	21.44	91.07	51.40		256 ALA		
N	31.50	87.73	11.96	CD	19.42	93.05	52.66	CE	57.26	79.11	88.33
				CD	21.48	94.49	52.42	CA	57.32	76.04	88.17
	238 GLU			CG	22.23	94.20	53.17	C	59.39	74.74	89.33
OE2	40.76	94.09	15.93	CB	25.09	93.93	53.42	O	60.47	75.23	91.72
OE1	39.08	97.65	16.99	CA	23.04	91.72	52.00	N	55.11	75.13	89.38
CD	38.94	95.72	16.67	C	24.48	89.69	53.65				
CG	38.53	93.44	18.24	O	26.33	89.43	51.67		257 SER		
CA	37.20	90.86	19.21	N	27.58	88.93	55.82	OG	66.30	72.56	89.31
C	37.40	89.66	22.70					CB	65.28	70.95	86.14
O	36.33	90.73	24.82					CA	64.11	72.03	88.17
N	34.80	89.57	18.04					C	66.13	74.74	89.14
								O	66.57	74.33	89.11
								N	61.54	73.79	87.54
									258 PHE		
								CD2	65.28	83.27	84.27
								CE2	62.41	85.85	87.32
								CZ	61.36	85.96	85.31
								CE1	61.74	85.47	84.32
								CD1	63.82	84.84	83.41
								CG	65.48	82.75	87.04
								CB	67.31	80.93	88.12
								CA	66.81	79.38	90.61
								C	69.61	81.36	92.01
								O	69.18	81.84	93.03
								N	65.39	77.12	89.83

Table 4 Conformational angles in HCAC.

NO.	NAME	PHI	PSI	OMEGA
3	HIS	-57.37	179.51	-71.44
4	TRP	-90.03	93.80	-123.03
5	GLY	-69.75	163.59	-179.03
6	TYR	-120.73	80.03	-171.05
7	GLY	-74.24	153.45	-179.04
8	LYS	54.20	48.52	-174.87
9	HIS	-122.74	94.03	172.31
10	PRO	JUD	JUD	JUD
11	GLY	-69.53	177.12	
12	PRO	JUD	JUD	
13	GLU	-75.44	153.73	-179.75
14	HIS	-97.30	147.39	-170.27
15	TRP	-70.73	-24.04	-177.27
16	HIS	-93.33	94.33	179.03
17	PRO	-93.32	147.04	-166.79
18	ASP	-74.03	19.39	-178.80
19	PHE	-138.91	94.01	-174.85
20	PRO	JUD	JUD	-75.39
21	ILE	-74.32	144.74	-179.09
22	ALA	-74.53	149.43	-178.03
23	LYS	-111.44	131.51	32.02
24	GLY	-74.03	153.28	-179.04
25	ARG	-74.01	147.44	179.03
26	ARG	-111.33	75.65	177.52
27	GLN	-74.03	144.74	-59.43
28	SER	122.42	144.37	179.09
30	VAL	-57.37	129.20	179.03
31	ASP	-111.44	134.64	-178.84
32	ILE	-111.33	117.83	47.37
33	ASP	-111.44	111.56	-179.44
34	THR	-74.01	144.74	-35.43
35	HIS	-74.01	144.74	-99.00
36	THR	-94.62	111.74	21.40
37	ALA	-74.04	144.74	
38	LYS	-92.27	147.03	177.74
39	TYR	-111.44	139.41	29.48
40	ASP	-91.72	111.56	168.74
41	PRO	JUD	JUD	JUD
42	SER	111.47	111.74	-67.79
43	LEU	69.23	119.74	179.57
44	LYS	91.14	111.74	-71.50
45	PRO	JUD	JUD	JUD
46	LEU	-111.44	123.21	111.74
47	PRO	111.44	104.12	111.74
48	VAL	129.44	111.74	171.27
49	SER	91.33	111.74	-50.13
50	TYR	-127.15	126.52	-9.51
51	ASP	-111.44	114.24	99.11
52	GLN	69.23	93.23	45.48
53	ALA	111.47	111.74	111.74
54	THR	-138.48	111.12	99.61
55	SER	-10.74	122.42	105.34
56	LEU	-69.23	111.74	179.04
57	ARG	147.17	111.74	-109.93
58	ILE	-134.74	109.52	140.56
59	LEU	-172.34	126.69	140.48
60	ASN	-140.53	113.84	173.57
61	ASP	-107.93	121.67	8.74
62	GLY	85.74	119.42	-3.89
63	HIS	-140.11	108.17	2.67
64	ALA	177.20	146.98	101.11
65	PHE	-100.54	99.00	132.48
66	ASN	-127.33	116.83	155.63
67	VAL	-138.75	106.87	147.64
68	GLU	-132.47	117.93	134.02
69	PHE	-107.93	105.10	169.48
70	ASP	-64.08	113.73	151.64
71	ASP	-117.69	138.89	-12.23
72	SER	-92.10	121.08	68.00
73	GLU	170.43	108.37	-101.37
74	ASP	-139.27	113.46	39.26
75	LYS	-49.08	120.68	8.38
76	ALA	-149.10	88.23	85.58
77	VAL	-133.41	123.52	179.36
78	LEU	-143.82	117.81	134.46
79	LYS	-136.91	103.63	170.91
80	GLY	153.38	110.45	151.28
81	GLY	-35.12	115.84	-50.36
82	PRO	JUD	JUD	-4.96
83	LEU	-143.07	120.16	134.05
84	ASP	-143.45	103.04	104.11
85	GLY	111.34	114.31	107.65
86	THR	-76.11	114.59	126.06
87	TYR	-57.37	71.77	156.27
88	ARG	-106.10	110.16	123.90
89	LEU	-109.94	112.08	138.48
90	ILE	-117.84	114.55	-53.45
91	GLN	-143.75	94.45	103.98
92	PHE	-144.1	104.27	151.80
93	HIS	-151.47	123.46	-156.62
94	PHE	-143.34	114.00	132.86
95	HIS	-123.55	115.32	128.90
96	THR	-142.23	126.38	-103.96
97	GLY	-131.99	120.14	-155.83
98	SER	-98.2	113.95	94.50
99	LEU	117.35	107.55	175.51
100	GLY	71.71	109.1	22.88
101	GLY	91.1	116.75	23.32
102	GLU	171.54	147.56	161.30
103	GLY	143.17	109.11	15.74
104	SER	-83.76	130.09	162.29
105	GLN	-111.55	97.53	73.87
106	HIS	97.46	109.96	156.14
107	PHE	137.33	109.1	168.41
108	VAL	133.21	108.5	161.1
109	ASP	71.45	113.57	94.83
110	LYS	84.99	109.21	14.04
111	LYS	111.47	111.42	145.35
112	LYS	138.12	123.80	151.84
113	TRP	133.04	110.71	155.83
114	ALA	-72.55	123.75	86.44
115	ALA	-152.60	108.64	170.70
116	GLY	83.54	108.76	112.77
117	LEU	56.71	111.33	135.83
118	HIS	119.57	109.24	114.56
119	LEU	-97.21	119.73	95.56
120	VAL	111.71	109.41	121.09
121	HIS	-125.61	98.75	160.78
122	TRP	111.44	147.55	98.52
123	ASN	-97.75	108.29	117.52
124	THR	92.7	121.06	-93.01
125	GLY	97.93	111.49	90.11
126	TYR	-130.42	115.75	9.62
127	GLY	73.03	122.79	-93.05
128	ASP	134.39	111.93	176.10
129	PHE	73.49	113.98	-90.20

65	PHE	-100.54	99.00	132.48	127.48	93.41			
66	ASN	-127.33	116.83	155.63	-128.78	45.29			
67	VAL	-138.75	106.87	147.64	89.25				
68	GLU	-132.47	117.93	134.02	-55.01	128.08	15.52		
69	PHE	-107.93	105.10	169.48	-67.41	88.81			
70	ASP	-64.08	113.73	151.64	-175.68	-50.04			
71	ASP	-117.69	138.89	-12.23	-113.39	150.79			
72	SER	-92.10	121.08	68.00	-8.39				
73	GLU	170.43	108.37	-101.37	-132.15	27.36	75.48		
74	ASP	-139.27	113.46	39.26	-117.31	18.98			
75	LYS	-49.08	120.68	8.38	72.50	-108.42	-145.07	115.35	
76	ALA	-149.10	88.23	85.58					
77	VAL	-133.41	123.52	179.36	-55.54				
78	LEU	-143.82	117.81	134.46	100.44	97.08			
79	LYS	-136.91	103.63	170.91	-26.78	160.77	125.35	179.84	
80	GLY	153.38	110.45	151.28					
81	GLY	-35.12	115.84	-50.36					
82	PRO	JUD	JUD	-4.96					
83	LEU	-143.07	120.16	134.05	-75.85	-169.27			
84	ASP	-143.45	103.04	104.11	51.49	75.35			
85	GLY	111.34	114.31	107.65					
86	THR	-76.11	114.59	126.06	-72.42				
87	TYR	-57.37	71.77	156.27	-132.31	-112.44			
88	ARG	-106.10	110.16	123.90	-123.31	176.95	-159.76	164.84	-6.53
89	LEU	-109.94	112.08	138.48	145.00	104.38			
90	ILE	-117.84	114.55	-53.45	-170.74	178.57			
91	GLN	-143.75	94.45	103.98	127.99	-67.89	93.12		
92	PHE	-144.1	104.27	151.80	80.69	98.38			
93	HIS	-151.47	123.46	-156.62	94.80	-90.67			
94	PHE	-143.34	114.00	132.86	-96.76	71.94			
95	HIS	-123.55	115.32	128.90	-74.24	-108.32			
96	THR	-142.23	126.38	-103.96	59.53	126.08			
97	GLY	-131.99	120.14	-155.83					
98	SER	-98.2	113.95	94.50	72.58				
99	LEU	117.35	107.55	175.51	-68.11	92.34			
100	GLY	71.71	109.1	22.88	76.84	113.47			
101	GLY	91.1	116.75	23.32					
102	GLU	171.54	147.56	161.30	40.92	133.71	-174.15		
103	GLY	143.17	109.11	15.74					
104	SER	-83.76	130.09	162.29	-130.84				
105	GLN	-111.55	97.53	73.87	74.04	161.37	-91.85		
106	HIS	97.46	109.96	156.14	71.62	-91.27			
107	PHE	137.33	109.1	168.41	28.77				
108	VAL	133.21	108.5	161.1	170.74				
109	ASP	71.45	113.57	94.83	74.27	-145.83			
110	LYS	84.99	109.21	14.04	-120.61	64.48	1...81	171.84	
111	LYS	111.47	111.42	145.35	-59.06	76.16	96.27	-114.45	
112	LYS	138.12	123.80	151.84	-69.72	145.07	155.48	80.43	
113	TRP	133.04	110.71	155.83	-74.97	83.15			
114	ALA	-72.55	123.75	86.44					
115	ALA	-152.60	108.64	170.70					
116	GLY	83.54	108.76	112.77	152.66	176.33	-83.46		
117	LEU	56.71	111.33	135.83	103.74	94.44			
118	HIS	119.57	109.24	114.56	98.99	169.05			
119	LEU	-97.21	119.73	95.56	-75.12	0.81			
120	VAL	111.71	109.41	121.09	171.61				
121	HIS	-125.61	98.75	160.78	-27.93	132.67			
122	TRP	111.44	147.55	98.52	87.87				
123	ASN	-97.75	108.29	117.52	-109.80	7.91			
124	THR	92.7	121.06	-93.01	-163.01				
125	GLY	97.93	111.49	90.11	87.12	134.13	-148.58	130.31	
126	TYR	-130.42	115.75	9.62	-81.99	164.41			
127	GLY	73.03	122.79	-93.05					
128	ASP	134.39	111.93	176.10	90.42	73.01			
129	PHE	73.49	113.98	-90.20	-166.98	-123.98			

130 GLY	-59.20	109.62	-25.42					
131 LYS	-90.07	125.74	-18.65	-94.14	79.39	128.57	149.06	
132 ALA	-108.79	111.24	-5.69					
133 VAL	-76.71	107.02	18.85	51.53				
134 GLN	-98.12	117.66	-20.08	-55.99	138.54	178.87		
135 GLN	-117.37	118.44	-178.86	-62.46	177.12	-130.09		
136 PRO	0.0	0.0	-60.64					
137 ASP	-99.95	129.23	50.40	90.04	92.38			
138 GLY	-84.58	111.43	-40.04					
139 LEU	-108.71	127.15	164.57	-101.59	47.31			
140 ALA	-139.53	107.49	119.40					
141 VAL	-115.17	139.65	122.23	173.97				
142 LEU	111.62	106.98	113.75	-165.70	-148.96			
143 GLY	-112.38	113.51	159.96					
144 ILE	-140.32	110.31	115.62	-22.50	-133.11			
145 PHE	-87.11	164.34	116.57	-81.69	-53.05			
146 LEU	-122.34	104.52	176.07	52.10	141.15			
147 LYS	-129.20	111.37	136.18	79.12	-140.87	-51.29	-125.04	
148 VAL	-23.25	108.30	124.10	-174.16				
149 GLY	-146.91	114.40	-121.92					
150 SER	-109.09	102.46	176.48	-69.04				
151 ALA	-56.58	110.36	151.82					
152 LYS	-133.70	91.03	89.98	-161.41	39.30	73.17	-14.43	
153 PRO	0.0	0.0	-63.83					
154 GLY	-67.08	113.35	5.01					
155 LEU	-116.35	111.40	-44.53	-138.04	94.96			
156 GLN	-48.25	112.27	-51.47	-136.05	12.15	24.07		
157 LYS	-49.76	115.38	-11.58	-41.46	6.93	-155.48	-141.16	
158 VAL	-57.37	72.49	-54.48	158.94				
159 VAL	-73.21	106.11	-71.15	161.00				
160 ASP	-24.87	114.42	-58.05	-90.79	-6.90			
161 VAL	-44.20	115.14	21.60	92.02				
162 LEU	-47.00	112.10	-6.84	-33.60	-178.65			
163 ASP	-55.48	104.14	-77.76	-111.14	-16.69			
164 SER	-39.47	122.30	-48.39	-26.32				
165 ILE	-81.37	127.33	20.48	-67.12	68.49			
166 LYS	-28.90	119.18	-44.44	-108.24	45.74	172.23	-144.85	
167 THR	-118.34	116.19	179.32	92.42				
168 LYS	-66.64	120.28	74.68	164.66	-107.26	-158.32	-69.56	
169 GLY	178.02	120.18	-57.88					
170 LYS	-52.25	118.74	141.13	-18.43	-125.22	-157.29	-177.18	
171 SER	-130.27	110.66	147.45	-25.47				
172 ALA	-128.03	110.98	144.63					
173 ASP	-69.43	108.31	117.60	-48.46	99.74			
174 PHE	-111.88	111.70	74.08	-150.85	71.18			
175 THR	-124.78	124.15	98.00	-50.25				
176 ASN	123.35	129.08	71.44	-62.94	125.13			
177 PHE	-97.92	137.01	-116.02	-130.11	51.13			
178 ASP	141.36	97.32	68.12	-139.01	151.01			
179 PRO	0.0	0.0	26.53					
180 ARG	-98.02	121.54	-43.03	-83.42	97.39	158.50	-59.13	3.17
181 GLY	-87.26	100.53	15.87					
182 LEU	-116.14	108.92	7.59	-141.96	-117.33			
183 LEU	-113.80	120.21	138.90	-97.91	103.96			
184 PRO	0.0	0.0	-178.16					
185 GLU	-108.83	113.75	-3.14	-175.33	-101.76	5.15		
186 SER	-123.23	100.80	132.65	149.92				
187 LEU	-86.19	111.80	-7.62	-91.00	-35.20			
188 ASP	-93.26	110.95	153.46	-69.23	169.72			
189 TYR	-156.58	127.01	162.80	59.85	80.85			
190 TRY	-111.44	104.76	145.79	-78.86	-93.79			
191 THR	-141.84	109.96	140.01	-156.60				
192 TYR	-179.38	113.32	151.23	92.19	96.15			
193 PRO	0.0	0.0	132.71					
194 GLY	-138.13	118.36	-168.00					

195 SER	-141.76	110.18	170.61	-144.47				
196 LEU	-79.79	91.80	155.40	-144.24	32.04			
197 THR	-103.37	115.61	6.92	38.39				
198 THR	-154.62	102.77	159.00	8.35				
201 LEU	-57.37	102.36	120.72	-58.17	40.52			
202 LEU	-37.84	109.50	170.12	-52.16	-17.76			
203 GLU	-129.56	113.36	50.25	-25.14	-154.94	103.56		
204 CYM	-108.74	128.92	59.10	49.55				
205 VAL	-132.31	120.31	153.28	15.87				
206 THR	-113.67	105.38	133.42	-78.04				
207 TRY	-101.41	120.04	111.63	-64.27	99.23			
208 ILE	-106.22	106.94	119.45	-19.93	-96.05			
209 VAL	-125.81	104.77	152.66	-177.74				
210 LEU	-117.77	113.63	120.59	-90.39	117.54			
211 LYS	-60.92	117.36	-19.97	113.09	-169.61	-65.44	-135.37	
212 GLU	-99.74	107.97	74.41	-135.07	-138.09	144.12		
213 PRO	0.0	0.0	173.00					
214 ILE	-113.08	102.23	158.29	56.11	162.28			
215 SER	-118.50	114.17	151.20	79.07				
216 VAL	-142.23	123.92	150.64	-23.15				
217 SER	-57.17	121.33	93.89	-115.78				
218 SER	-2.65	103.94	-22.85	165.33				
219 GLU	-69.11	140.15	-1.28	-69.64	-19.23	-46.13		
220 GLN	-121.26	84.47	-26.06	-96.60	137.93	-18.30		
221 VAL	-56.46	104.32	-69.40	120.25				
222 LEU	-43.96	120.67	-93.66	-62.59	-140.42			
223 LYS	-61.58	106.27	-65.26	-36.02	-121.07	54.75	69.75	
224 PHE	-19.79	111.65	-00.86	-83.76	31.17			
225 ARG	-53.54	122.36	14.15	-37.73	-154.13	156.28	105.58	-21.61
226 LYS	-101.81	108.49	-11.05	-104.02	147.44	-165.15	-12.09	
227 LEU	-58.49	111.95	165.26	-93.57	-6.30			
228 ASN	-127.83	112.41	130.38	-47.73	73.78			
229 PHE	-02.83	116.97	-37.78	-80.36	137.59			
230 ASP	-137.69	108.88	-160.43	24.26	76.27			
231 GLY	-118.44	122.24	-126.44					
232 GLU	-90.93	107.23	-1.07	-88.35	35.97	110.93		
233 GLY	-122.79	104.81	44.46					
234 GLU	178.52	108.13	85.39	-107.29	4.26	-158.62		
235 PRO	0.0	0.0	-111.15					
236 GLU	3.82	117.44	-76.55	1.03	-55.67	-164.08		
237 GLU	13.95	120.49	68.91	71.17	-89.34	-79.07		
238 LEU	-93.43	109.35	86.79	-117.05	-122.60			
239 MET	-83.49	112.26	91.92	166.42	173.03	-163.50		
240 VAL	-142.33	125.03	174.71	-82.92				
241 ASP	38.39	116.04	54.38	-44.19	134.90			
242 ASN	-101.25	116.16	55.92	55.57	-176.34			
243 TRY	-112.31	116.82	-163.12	52.36	-91.53			
244 ARG	-140.84	113.89	121.37	-135.37	-137.99	-179.63	157.08	-175.52
245 PRO	0.0	0.0	147.16					
246 ALA	-55.49	107.51	150.14					
247 GLN	-105.58	112.88	157.88	-86.42	-52.13	-94.31		
248 PRO	0.0	0.0	164.28					
249 LEU	-97.95	126.24	-55.70	-111.14	-30.37			
250 LYS	60.93	108.46	-147.03	-126.86	13.88	98.30	-111.85	
251 ASN	-57.56	140.28	43.87	-107.96	-170.55			
252 ARG	-141.34	136.47	138.58	-43.64	160.47	-90.66	176.75	-167.39
253 GLN	-97.01	115.43	149.21	-111.85	-130.88	-33.10		
254 ILE	-120.40	117.91	156.17	-153.39	-176.00			
255 LYS	-133.70	111.12	163.92	-32.55	99.03	161.46	-69.70	
256 ALA	-137.51	108.53	131.81					
257 SER	-70.36	111.71	21.51	73.21				
258 PHE	-150.13	116.22	147.53	77.79	-93.40			

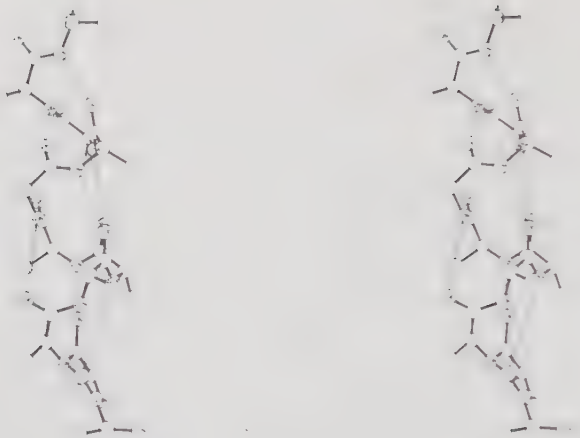


Fig. 15 An ORTEP stereo drawing of the G-helix.

SECONDARY AND TERTIARY STRUCTURE

Helical segments

The HCAC molecule contains 7 right-handed helical segments (labelled A to G) listed in table 5 . The main chain conformation is considered helical when a segment contains at least one hydrogen bond, and a hydrogen bond has been assumed to be present when the carbonyl oxygen and the amide nitrogen of the main chain are within 2.6-3.3 Å and the O---H-N angle is between 140 - 180 degrees.

Table 5. Secondary structures in HCAC.

HELIX	3_{10}-BEND TYPE I OR II	β-STRUCTURE	CHAIN NUMBER IN THE PLEATED SHEET	AMINO ACIDS INVOLVED
	I			7 GLY- 10 ASP
	I			12 PRO- 15 TRP
A_{α}				15 TRP- 19 PHE
$B_{3/10}$				20 PRO- 24 GLY
C_{α}				33 ASP- 37 ALA
		X	10	38 LYS- 40 ASP
		X	2	54 THR- 61 ASN
		X	3	64 ALA- 70 ASP
	I			80 GLY- 83 LEU
		X	4	88 ARG- 95 HIS
		X	5	115 ALA-123 ASN
	I			123 ASN-126 TYR
D_{α} 3/10				128 ASP-135 GLN
	I			135 GLN-138 GLY
		X	6	138 GLY-149 GLY
E, irregular				153 PRO-166 LYS
		X	1	169 GLY-172 ALA
$F_{3/10}$				179 PRO-183 LEU
		X	8	189 TYR-194 GLY
		X	7	204 CYS-210 LEU
G_{α} 3/10				217 SER-227 LEU
		X	9	254 ILE-257 SER

Table 6. Other β-structures; Segments A_1 and A_2 etc. hydrogen bonded colinearly to each other

SEGMENTS	AMINO ACIDS INVOLVED
A 1	29 PRO
A 2	247 GLU
B 1	30 VAL - 32 ILE
B 2	107 THR - 109 ASP
C 1	44 LYS - 52 GLN
C 2	81 GLY - 75 LYS
D 1	86 THR - 89 LEU
D 2	78 LEU - 76 ALA
E 1	86 THR - 87 TYR
E 2	124 THR - 122 TRP
F 1	97 Gly
F 2	102 GLU
G 1	108 VAL
G 2	111 LYS
H 1	213 PRO - 216 VAL
H 2	145 PHE - 148 VAL
I 1	227 LEU - 229 PHE
I 2	239 MET - 237 GLU

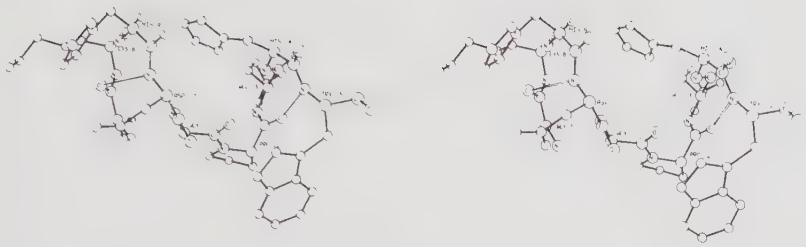


Fig. 16 An ORTEP stereo drawing of the 3_{10}-bends at residues 7-10 and 12-15.

The A helix is of the α-type and the B, and F helices are of the 3_{10}-type. Helix A contains one hydrogen bond whereas the B, and F helices contain two. The D and G (fig. 15) helices begin in the α-helix conformation and end in a 3_{10}-conformation. The longest helix, E, is a mixture of 3_{10}- and α-helix conformations and is very irregular (1). This helix ends in α_{II}-conformation. In the helices B and E there are proline residues in the amino termini. A proline residue has been found at the amino termini of helices in other proteins (42). In HCAC a proline residue also appears as the first residue beyond the C-terminal residue of the A, D, and F helices.

Turns

In several proteins the main chain has been found to fold back upon itself by forming a hydrogen bond between the carbonyl group of residue n and the amide group of residue n + 3. These hairpin loops are of two different types, I and II, as described by Venkatachalam (43).

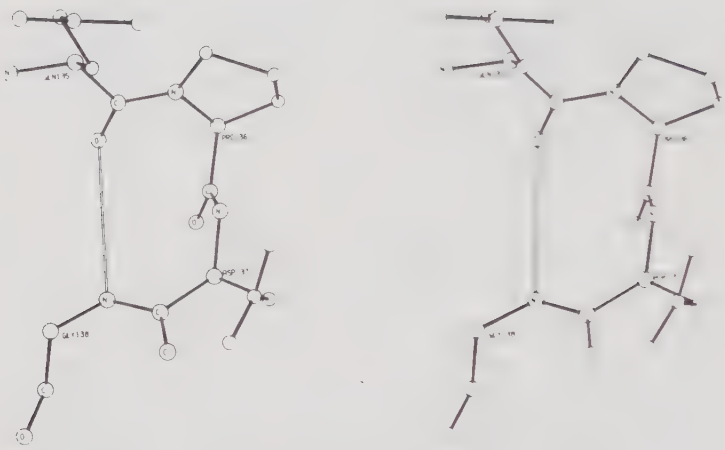


Fig. 17 An ORTEP stereo drawing of the 3_{10}-bend at residues 135-138.

There are many reverse turns in the molecule. Most often they appear when a chain of the large pleated sheet folds back. They are mainly located on the surface. 3_{10}-bends are formed between residues 7-10, 12-15, 80-83, 123-126 and 135-138 and are all of the type I (43). Fig. 16, 17 show the 3_{10}-bends at residues 7-10 and 12-15 and the 3_{10}-bend at residues 135-138, respectively (table 5).

The 3_{10}-bend at residues 80-83 is very similar to that of 130-135 and the one at 123-126 ressembles that at 7-10. The backbone at residues 7-15, including two 3_{10}-bends, has a nearly helical conformation, but Pro 12 prevents helix formation by blocking the formation of the required hydrogen bond.

β-structures

A large ten-stranded β-structure runs through the molecule. It has been described earlier (1, 2). The strands are antiparallel except

for two pairs, strands 6-7 and 9-10. The main chain of the β-structure is not completely extended in some places, which causes a slightly irregular hydrogen bonding pattern (fig. 18), thus not all of the residues in the sheet form main chain hydrogen bonds. The β-structure has a right-handed twist of 220 degrees from the left to the right edge (1, 2). Right-handed twists of β-structures have been found in several other proteins, which contain large β-structures. Although in carboxypeptidase eight segments of the peptide form a similar twisted pleated sheet, the sequence of the chain segments is different from that in HCAC.

In general β-structures are found in the interior of the proteins with only the edges exposed to solvent. In this way mainly hydrophobic residues are involved in the formation of these β-structures. However, in carbonic anhydrase which has a very large active site opening, (fig. 19) all peptide segments except 8, 7 and part of 6 are exposed to solvent. During the chemical determination of the amino acid sequence great difficulties were encountered in sequencing a peptide consisting of 60 amino acids (44). This peptide includes strands 4, 5 and 6 of the large pleated sheet. The difficulty in digesting that peptide might have been due to the tight hydrogen bond arrangement (fig. 18) and the large number of hydrophobi residues. The sequence was eventually solved by methods described b· L. Strid et al. (45). The ten-stranded pleated sheet contains about 52 residues or 20 % of the residues in HCAC. There are other β-structures on the surface of the molecule. They often appear when chain segments in the large pleated sheet make reverse turns, and therefore, most of them are antiparallel (table 6).

In the HCAC molecule 35 % of the residues are in pleated sheet, 12 % in 3_{10}-bends and 20 % in helices. In a study on the amount of secondary structure in eleven proteins Crawford et al. (42) found 17 % of the residues in β-sheet, 32 % in turns of all types and 34 % in helices. These figures indicate the importance of secondary struc-

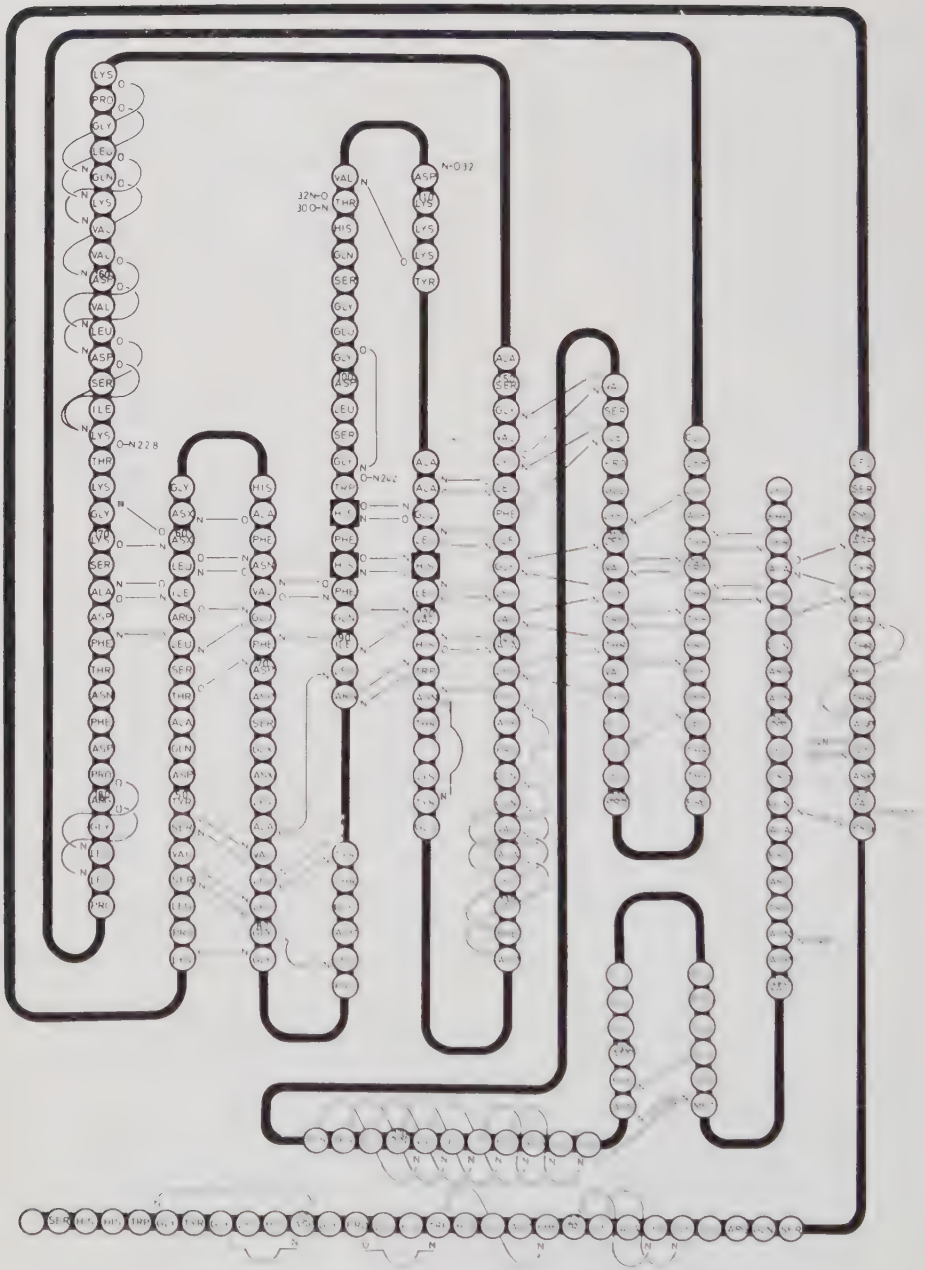


Fig. 18 The main chain hydrogen bonds in HCAC. The zinc ligands are in squares. The β -structure is the large hydrogen bond system in the center of the drawing. The helices are indicated by the curved lines and the $3/_{10}$-bends by the bows.

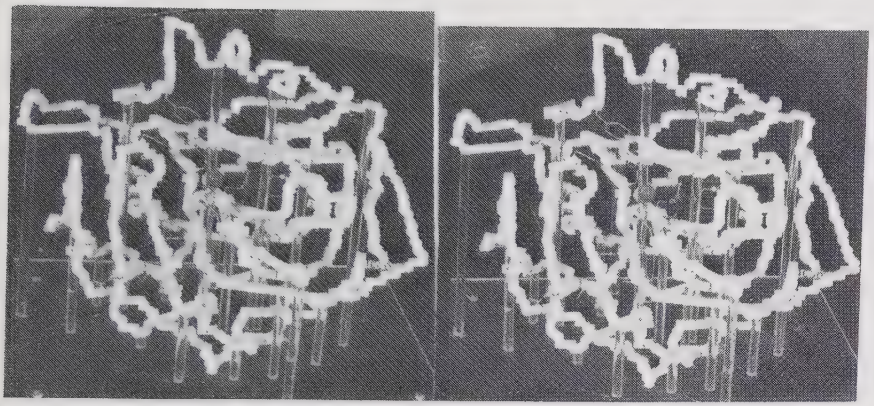


Fig. 19 Stereoscopic photograph of the main chain folding

tures as conformational units when predicting 3-dimensional structures from known sequences, as most of the amino acids in proteins seem to be involved in such structures.

Aspects of folding

Carbonic anhydrase cannot be folded sequentially from the amino end to the carboxylic end, since segments of the polypeptide chain synthesized later are folded between segments synthesized earlier. (fig. 19, 21). Four interesting folding experiments involving denaturation and renaturation of HCAC have been reported.

1) The zinc-containing enzyme was denatured with 6 M guanidinium hydrochloride and renatured by dialysis against low salt buffer yielding products with almost 100 % of the specific activity of the native enzyme (46).

2) Zinc was removed from the enzyme by dialysis with EDTA, and then the enzyme was denatured. Zinc was added to the denatured apoenzyme solution and then the enzyme was renatured by dialysis as above. The CD and ORD spectra indicated that refolding was rapid. The product was active. Similar results have been obtained with BCAB (47).

3) The apoenzyme was denatured and renatured by dialysis as described above without zinc addition. The product was inactive but when zinc was added full activity was restored. The folding of the apoenzyme was slower than that of the zinc containing enzyme (48).

4) If the apoenzyme was partially renatured, the zinc added, and renaturation completed, the product showed no substrate activity. Further, spectral studies showed that the structure of this "renatured enzyme" was different from that of the products of the first three experiments (48).

The experiment which yielded an inactive enzyme may be explained, if the zinc ions bind to secondary sites which are available only in a partially folded enzyme, and that this binding prevents formation of the "native enzyme" structure.

Location of various side chains

The complete sequence of HCAC is known and thus the various side chains can be located in the model. The side chains in the active site will be discussed later.

Two of the five arginyl residues are exposed. All lysyl residues extend out into the solvent, though some are involved in hydrogen bonds. The accessability of these basic side chains agrees with

hydrogen ion equilibria studies on HCAC (50) as well as with chemical modification studies of the lysyl residues in BCAB by Nilsson et al. (51), and in HCAB by Whitney et al. (52). Only one charged residue, Glu 116, is located in the interior of HCAC. This residue is near the active site and will be discussed in the next chapter.

The NMR studies of Campbell et al. (53) indicated that there are 6 titrateable imidazoles in HCAB, of which three are sensitive to the presence of inhibitors and to modifications made in the active site. Other NMR studies have yielded the pK values for the titrateable histidyl residues in HCAC (54). The structure of HCAC suggests that there are 7 exposed histidyl residues of which only one, His 63, is located in the active site (table 8). Riddiford et al. (55) have found that five of the twelve histidyl residues in HCAC titrate normally.

All phenylalanyl residues are located in the interior of the molecule, except for Phe 19, which is in an aromatic cluster (table 7) but is partly exposed and Phe 129 which is located in the opening of the active site and is exposed. Four of the eight tyrosyl residues are exposed. Of these Tyr 87 is only partly exposed and Tyr 39 is exposed but in the crystal hydrogen bonded to Asp 137 in a neighbouring molecule. This agrees well with the titration results of Riddiford et al. (56) and the results of nitration of tyrosyl residues in HCAC (57). All seven tryptophanyl residues are unambigously located. Four of them (4, 122, 207 and 243) are partly exposed to solvent and three (15, 96 and 190) are in the interior of the molecule. This agrees with acid difference spectra of Riddiford et al. (58). The single cysteinyl residue in HCAC is buried. However, there is no steric hindrance to rotation of the SH-group about the C_α - C_β bond, and thus the SH group can be accessible to the solvent region which is necessary for the observed labeling with mercurials.

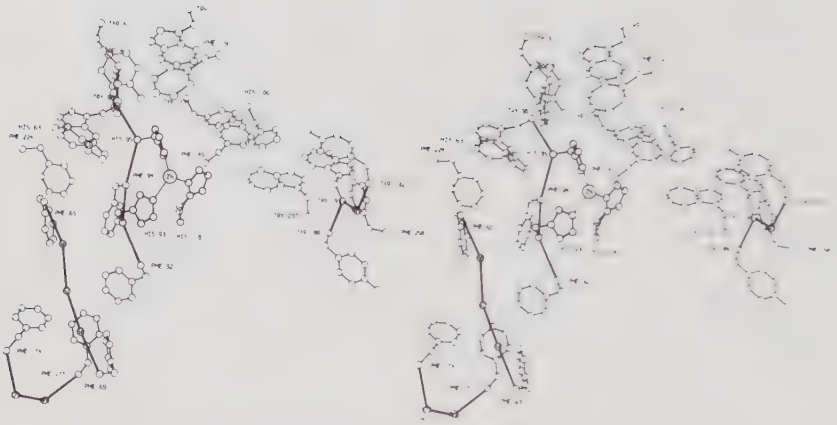


Fig. 20 Aromatic clusters in HCAC

Table 7 Aromatic clusters.

I		II		III	
Residues in HCAC	In HCAB if different	Residues in HCAC	In HCAB if different	Residues in HCAC	In HCAB if different
PHE 65		HIS 3	ASP	HIS 106	
PHE 69		TRP 4		TYR 113	
PHE 92		TYR 6		PHE 145	LEU
PHE 94		TRP 15		TRP 190	
PHE 96		PHE 19	TYR	TYR 192	
PHE 175		HIS 63		TRP 207	
PHE 177				TYR 189	PHE
PHE 224				PHE 258	

Table 8. Exposed aromatic residues

Residue type	Exposed	Partially exposed
HIS	2, 3, 9, 14, 16, 35, 63	93, 95, 106, 118
PHE	129	19
TRP		4, 122, 243
TYR	87	39, 113, 126

Comparison with spectral studies

CD and ORD measurements have been performed on the two human carbonic anhydrases, and indicate a large content of β-structure (59). Timasheff, Susi and Stevens (62) predicted the same for bovine carbonic anhydrase using IR spectra.

The 240 to 310 nm region of the ORD and CD spectra is rich in detail. The multiple cotton effects observed are thought to be due to aromatic chromophores in specific asymmetric environments (63 , 64). In the structure of HCAC there are many aromatic residues near the surface, and there are three aromatic clusters (table 7, 8; fig. 20) which would contribute to the observed cotton effects.

However, at neutral pH the usual "helical" trough is obscured by the very pronounced aromatic cotton effects. These effects vanish on acid denaturation (59, 60). At pH 1.6-1.8 Beychook (59) has found that the amount of α-helix is about 20 %. Further difficulties in estimation the α-helix content are caused by the large pleated sheet and the irregular helices.

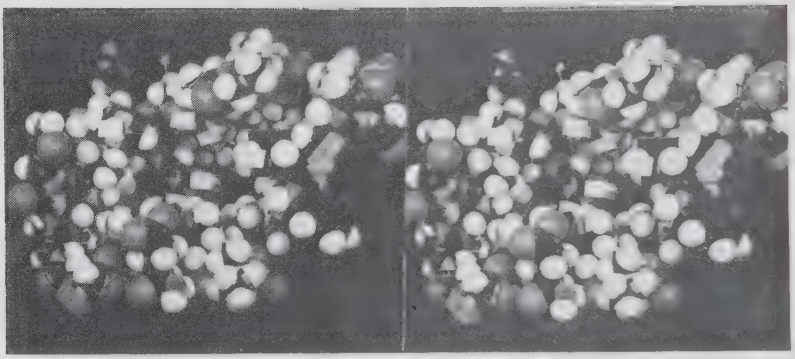


Fig. 21 Active site region of HCAC

ACTIVE SITE

The metal ion

The zinc ion was easily located in the electron density map, appearing as the highest density. The active site is a 15 Å deep cavity, at the bottom of which the essential zinc ion is situated. The roughly circular, 15 Å diameter opening of the cavity is defined by His 63, Pro 200, Phe 129 and the third strand in the pleated sheet (fig. 21). The zinc ion is bound to the protein via three histidyl residues of which His 93 and His 95 are linked to the zinc ion by their $N\varepsilon$ nitrogens, and His 118 is linked by its $N\delta$ nitrogen. His 93 and His 95 are located in strand 4 and His 118 in strand 5 of the large pleated sheet.

Table 9. Distances between the zinc ion and its ligands.

His 93 (Nε)	His 95 (Nε)	His 118 (Nδ)	H_2O
2.4 Å	2.0 Å	2.0 Å	1.9 Å

The imidazole planes of His 95 and His 118 are directed towards the center of the metal ion. This is not the case for His 93. Furthermore the electron density is larger for His 93 and the position of the imidazole ring varies between two extreme positions which differ by 0.7 Å. His 93 is also the most exposed ligand. The shortest zinc-ligand distance, involving His 93, is 2.4 Å. Table 9 gives all zinc-ligand distances. The difference in zinc-ligand distances as well as the position of the imidazole plane suggests that His 93 is bound differently to the zinc ion when compared to His 95 and His 118. His 93 shows the same apparent difference in the HCAB crystal structure. The zinc ion has a fourth ligand probably a water molecule or a hydroxyl ion giving the metal ion a somewhat distorted tetrahedral coordination. The largest deviation from tetrahedral angles is about 20°.

In insulin (65) the zinc ion has an octahedral coordination, provided by three Nε imidazole nitrogens from identical residues in three insulin molecules and by three water molecules. In thermolysin (66) and carboxypeptidase (5) the zinc ion has approximately tetrahedral coordination. Two of the ligands are Nδ nitrogens of histidyl residues, the third is a glutamic acid and the fourth a water molecule. The results from HCAC and these other zinc metallo-proteins agree well with investigations on complexes of imidazole and metal ions, although the studies in most cases showed that the Nδ atom was bound to the metal ion (67). Many experiments have been performed on the cobalt substituted carbonic anhydrase to study the metal ion coordination. Absorption spectra (68-71) and magnetic susceptibility experiments on cobalt BCAB (72) indicate a distorted tetrahedral coordination for

Table 10. Exposed side chains in the active site of HCAC and HCAB.

POLAR RESIDUES HCAC		HCAB IF DIFFERENT	NONPOLAR RESIDUES HCAC		HCAB IF DIFFERENT
TYR	6		ALA	64	SER
ASN	60		ILE	90	PHE
HIS	63		VAL	120	ALA
ASN	66	HIS	PHE	129	LEU
GLU	68	ASN	LEU	139	
GLN	91		VAL	141	
HIS	93		GLY	143	
HIS	95		PRO	199	
HIS	118		PRO	200	
THR	197		VAL	205	
THR	198	HIS	VAL	209	ILE

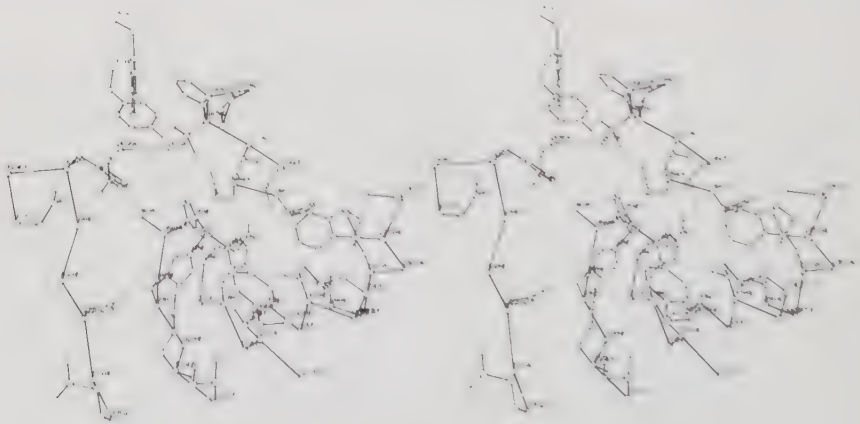


Fig. 22 An ORTEP stereo drawing of the active site region of HCAC

the metal ion. The interpretation of the visible absorption spectrum of cobalt HCAC (73, 74) suggests that at least at low pH the metal ion is bound to the enzyme via three nitrogen atoms and that these nitrogens could belong to histidyl residues. Electron paramagnetic resonance studies of the apoenzyme complexed with cobalt cyanide indicates an octahedral arrangement around the metal (75).

Electron spin resonance spectra on the copper substituted carbonic anhydrase (76) indicate that the metal is bound to two or three equivalent nitrogen ligands.

The zinc ion has been removed from the human C enzyme crystals by 2, 3-dimercaptopropanol-1 (BAL) (77). High resolution difference electron density maps of the $\underline{h}0\underline{1}$ and $0\underline{k}\underline{l}$ projections between the Zn(II), Co(II), Cu(II), Mn(II) and Hg(II) carbonic anhydrases and the apoenzyme indicate that there is no significant difference in the metal binding sites of the active metal carbonic anhydrases (Zn, Co) and the inactive (Cu, Mn) complexes. When the metal was replaced by mercuric chloride, the position of the metal was shifted toward Thr 197. In the isozyme HCAB, where the Hg(II) substituted enzyme was a heavy atom derivative, three-dimensional 2.0 Å data were available and the Hg was probably bound to Thr 197. The other ligands are His 94, His 96, His 119 (equivalent to His 93, His 95 and His 118 in HCAC) and two chloride ions giving the Hg an octahedral coordination.

Various residues in the active site

Table 10 lists the exposed amino acid residues in the active site cavity. The cavity contains approximately equal numbers of hydrophobic and hydrophilic residues. The residues are grouped such that the active site can be divided into a hydrophobic side and a hydrophilic side (fig. 22). Cys 204 is located near the hydrophobic region but is 12 Å from the zinc ion and shielded from the active site by Leu 196 and Leu 139.

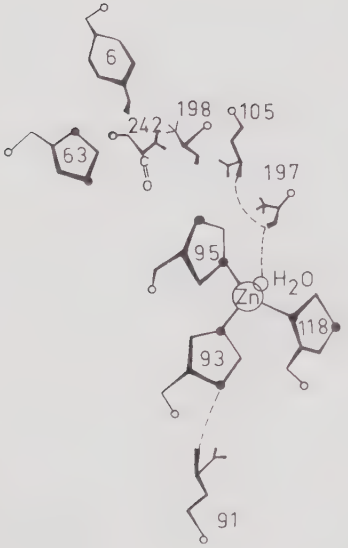


Fig. 23 A schematic drawing of some of the hydrogen bonds in the active site

The cysteinyl residue can be modified with a spin label either in native enzyme or denatured enzyme (which can be renatured). With the label attached to the sulfhydryl group, the hydration of carbon dioxide was not effected but the hydrolysis of esters was inhibited (78). The binding of esters might be effected either by conformational changes of Leu 196 and Leu 139 when the spin label binds to Cys 204 on the surface of the molecule, or by spin label binding to the cysteinyl residue from the inside of the active site facilitated by movement of Leu 196 and Leu 139.

In HCAB there is a serinyl residue in the position of the cysteinyl residue in HCAC, and the serinyl residue protrudes into the active site cavity. Thus the hydrophobic side in HCAC is less pronounced in HCAB.

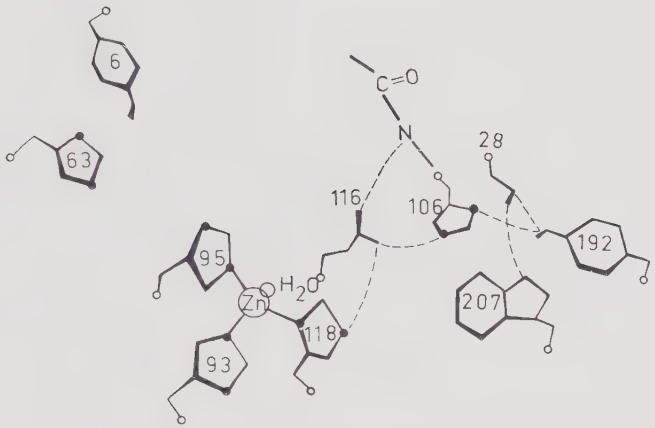


Fig. 24 A schematic drawing of some of the hydrogen bonds in the active site

There are a number of hydrogen bond interactions in the active site of HCAC. Thr 197 (Oγ) is 4.2 Å away from the zinc ion and is hydrogen bonded to the ligand water molecule (fig. 23). Thr 197 is also hydrogen bonded to Gln 105 (Oε). The zinc ligand of His 93 (Nδ) is hydrogen bonded to Gln 91 (Oε) (fig. 23). The Nδ group of His 95 is within hydrogen bonding distance of the peptide oxygen of Asn 242. However, the angle O----N-H is roughly 130 degrees and bonding is unlikely.

His 118 is involved in a complicated hydrogen bond system shown in figures 24 and 25. There are 7 residues involved in the system. His 118 (Nε) is hydrogen bonded to Glu 116 (Oε1) and the latter also to His 106 (Nδ). Glu 116 (Oε2) is also hydrogen bonded to the peptide nitrogen of His 106; His 106 (Nε) to Tyr 192 (Oη); Tyr 192 (Oη) to Ser 28 (Oγ); and Ser 28 (Oγ) to the peptide nitrogen of Ser 195 and to Thr 207 (Nε). This is a well defined region in the electron density map.

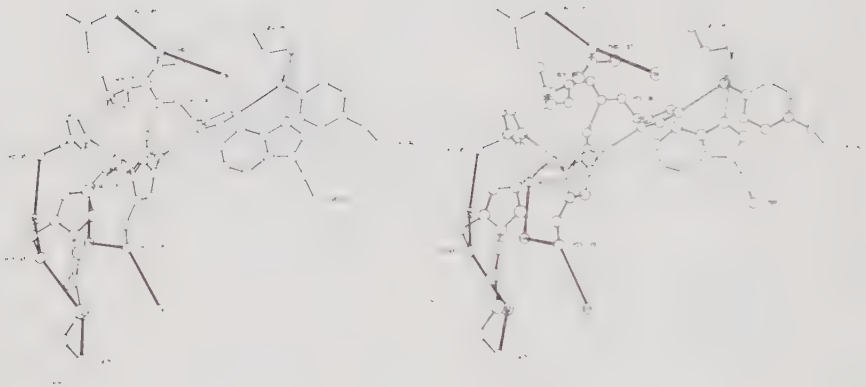


Fig. 25a An ORTEP stereo drawing of the hydrogen bonds in the active site of HCAC

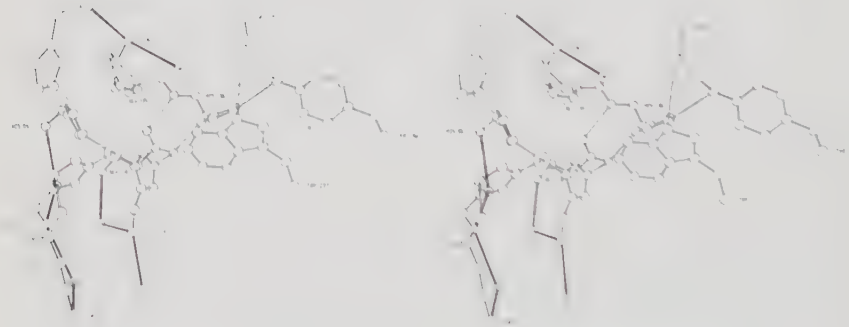


Fig. 25b An ORTEP stereo drawing of the hydrogen bonds in the active site of HCAB

His 106 is half buried with the Nδ atom more or less exposed. Tyr 192 is buried. Ser 28 and Trp 207 are half buried. However, the Nε atom of Trp 207 is buried. At this stage it is difficult to evaluate the importance of this system, however, HCAB has an identical system of hydrogen bonds around His 118, and, therefore it is possible that this hydrogen bond system is a "relay system" similar to that of the serinyl proteases and may be of importance in the catalytic mechanism. Furthermore, Glu 116 which is the only charged residue in the interior of the HCAB and HCAC molecules might be the "key" to the "relay system". The residues involved in this hydrogen bond network has been observed to be invariant in the primate, ox and sheep enzyme (60).

Another important group in the active site is His 63. Bromoacetazolamide and bromopyruvate are two inhibitors which covalently modify HCAC (79, 80). Göthe et al. have isolated a peptide containing a bromopyruvate labelled histidyl residue (81). The composition of that peptide agrees with the sequence around His 63. Kandel et al. have shown that bromoacetazolamide modifies His 63 (61). Table 11 gives distances between the zinc ion and certain atoms in the active site of HCAC.

The "water structure"

A possible "water structure" bound to the protein in the active site region has been described earlier (1). The earlier proposed link between the protein and the possible water structure must now be modified as residue 128, earlier assumed to be a histidyl residue, is a phenylalanyl residue. Further, the eight water molecules were placed in the same location as several of the heavy-atom groups, and it is possible that the densities interpreted as water molecules might be due to errors in phasing by the multiple isomorphous replacement method. A difference Fourier map based on the refined atomic coordinates of HCAC, which will be ready in the near future,

ought to give more reliable information on the possible active site water structure.

It is of interest to note that the nuclear magnetic relaxation rate of solvent water protons (82) can be interpreted in such a way that there is a water molecule or hydroxyl ion bound to the metal ion. It is often assumed that this liganded water molecule or hydroxyl ion participates in the catalytic mechanism (83, 84, 85-90). The assumption that the amino acids which are involved in the mechanism are the same in the two isoenzymes HCAB and HCAC is also reasonable. If this is correct, the only residues which can directly be involved in the mechanism are Thr 197, Gln 91, His 63 and the ligands to the metal ion.

Table 11. Distances between the zinc ion and certain atoms in the active site of HCAC

HIS	93	Nε	2.4 Å	LIGAND
HIS	95	Nε	2.0	"
HIS	118	Nδ	2.0	"
GLN	91	Oε1	6.9	
HIS	63	Nε	6.3	
TYR	6	Oη	6.8	
THR	198	Oγ	5.6	
THR	197	Oγ	4.2	
GLU	116	Cε2	6.6	
HIS	106	Nδ	8.1	
TYR	193	O	11.0	
TRP	207	Nε	9.6	
TRP	207	Cη	6.2	
GLU	68	Cδ	13.5	

ANION BINDING TO HCAC

The zinc metallo-enzyme carbonic anhydrase is inhibited by sulfonamides and anions (12, 10). The different carbonic anhydrase- anion complexes can be ordered by their dissociation constants:
$CN^- < SCN^- \simeq N_3^- < I^- < HCO_3^- < Br^- < CH_3COO^- < Cl^-$ (91).
This order of strength of interaction on proteins is called a Hofmeister lyotropic series (92). The order of binding strength for the halides is the reverse of what would be expected for binding to the free zinc ion. Anions inhibit the enzyme non-competitively with respect to substrates (93, 94, 95, 96). The binding of CN^-, SH^- and N_3^- cause significant changes in the absorption spectra of the Co(II) enzyme (97). Binding of chloride and bromide (97, 98) cause a pronounced line broadening in the NMR spectrum of Co(II) HCAC. In the case of chloride the pronounced line broadening is dependent of both cobolt and zinc ions and also on pH (99). Three of the heavy atom derivatives used in the structure investigation of HCAC (I^-, $Au(CN)_2^-$ and AMSulf) are inhibitors.

The inhibitors were located using two different types of difference electron density maps. The coefficients for these two are,

A) $(|F_H| - |F_p|)$, "single difference" or

B) $(2 |F_H| - |F_p|)$, "double difference"

where F_H and F_p denote the structure factors for the heavy atom derivative and the native protein, respectively. In the calculations the best phase angles obtained in native enzyme structure calculation were used.

The anion complexes studied in the crystal were HCAC-Cl^-, -Br^-, -I^-, -SCN^-, -$Au(CN)_2^-$, -$Pt(CN)_4^-$ and -$AuCl_4^-$ (table 12). For Cl^-, SCN^-, $Pt(CN)_4^-$ and $AuCl_4^-$ only projection data were used to locate the binding position (3).

Table 12. Location of various anions in HCAC

DERIVATIVE	RESOLUTION (Å)	SYMBOL	BINDING SITE IN THE PROTEIN
E-Cl^-	2.0 (projections only)	Cl_3	Zn, Thr 197
E-Br^-	2.5 (3-dimensional)	Br_3	Zn, Thr 197
		Br_6	Lys 23, Arg 244, Pro 12, Pro 245, Glu 13
E-I^-	2.5 (-"-)	I_3	Zn, Thr 197, Leu 196, Tyr 207
		I_6	Identical with Br_6
		$I_?$	Gln 91, Phe 129
E-SCN^-	2.0 (projections only)	SCN_3	Zn
E-$Au(CN)2^-$	2.5 (3-dimensional)	Cl_8	Zn, Gln 91
E-$Pt(CN)_4^-$	2.0 (projections only)	Pt_8	Gln 91, Phe 129, His 63
E-MMTGA +$Pt(CN)_4^-$	2.0 (-"-)	MPt_8	His 63, Zn, Cys 204
E-$AuCl_4^-$	2.0 (-"-)	Au_8	Gln 91, Phe 129
E-MMTGA + $AuCl_4^-$	2.0 (-"-)	MAu_8	His 63, Zn, Cys 204
E-MMTGA	2.0 (3-dimensional)	ML	Cys 204

In the HCAC-Cl^- and -SCN^- complexes the anion binds in the active site near the zinc ion (3) and within hydrogen bonding distance to Thr 197 ($O\gamma$). $Pt(CN)_4^-$ and $AuCl_4^-$ ions bind near Gln 91, however, if MMTGA modified enzyme crystals were used, the anion position is closer to the zinc ion and to His 63.

The three dimensional data for the HCAC-I^- complex were collected to 2.5 Å. A "double difference" map showed three major peaks in the active site, which were interpreted as a single I_3^- ion (3). However, the middle peak in the row of three also appears in the "double difference" maps of the HCAC-Diamox and HCAC- Br^- complexes. This site did not refine properly in any of the complexes nor did the peak appear in the "single difference" maps and, thus, it seems likely that the middle peak was caused by imperfect phasing. The outer peaks I_3 and I_7 (table 12) appear in the "single difference" maps and refined well. One iodide ion (I_3) is bound to Thr 197 ($O\gamma$) and to the NH group of the same residue and is in van der Waals contact with Leu 196 ($C\alpha$ and $C\delta$) and Trp 207. The other ion (I_7) is bound to Gln 91 and is in der van der Waals contact with Phe 129. Another iodide ion (I_6) is bound outside the active site cavity on a depression in the molecular surface (table 12).

The three-dimensional data for the HCAC-Br^- was collected to 2.8 Å resolution (table 12) and the "single difference" map indicates that bromide is bound in the same locations as I_3 and I_6 (table 12). The Br^- and I^- ions are located 3.0 Å and 3.4 Å from the zinc ion, respectively, and therefore it is unlikely that there is a solvent molecule between the anion and the zinc ion. The larger I^- ion (I_3) has more contact with the protein than the smaller Br^- ion which may explain why I^- is a more effective inhibitor than Br^- (3). Fig. 26 shows the binding of Br^- in the active site.

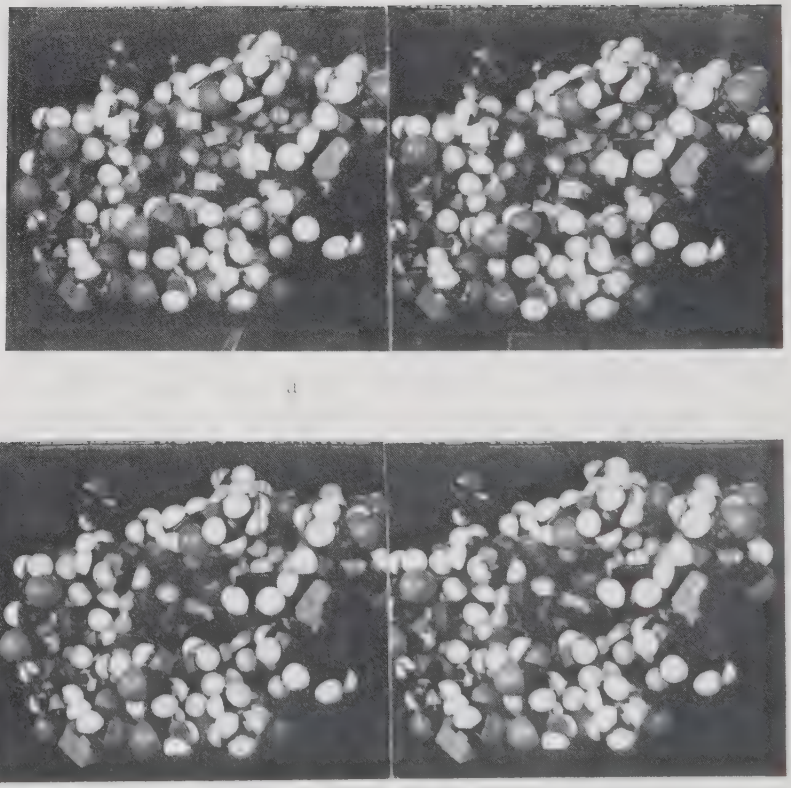


Fig. 26 Packing model of a) the active site of HCAC and b) the binding of Br^- in the active site

The three-dimensional data for the HCAC-$Au(CN)_2^-$ complex were collected to 2.5 Å. The linear gold cyanide ion binds to the metal bound solvent molecule and stretches across the active site to Gln 91 (4). A change of reactivity of His 63 has been detected in the crystal when studying the cooperativity in the binding of MMTGA and anions (4). This is confirmed with modification experiments of His 63 by bromopyruvate, which showed that the modification was accelerated in the presence of $Au(CN)_2^-$ in the active site (100).

In an attempt to locate the substrate binding site native enzyme crystals were soaked in 0.5 M $NaHCO_3$ solution in 0.05 M Tris-HCl buffer at pH 8.5 for at least six days. The HCO_3^- ion is a substrate at physiological pH and an inhibitor at higher pH values. The 2.5 Å three-dimensional data were collected for the HCAC-HCO_3^- complex. A few peaks were observed in the "single difference" map of which one was located in the active site near Tyr 6, Asn 242 and His 95 (metal ligand). This peak was 4.5 standard deviations above the background.

SULFONAMIDE BINDING TO HCAC

The inhibition of carbonic anhydrase by sulfonamides has been studied extensively (101, 102, 103, 104, 105). It has been found that the aromatic or heterocyclic sulfonamides are considerably stronger inhibitors than the aliphatic ones, and that substitutions on the ring cause marked changes in the inhibitory strength (106, 107).

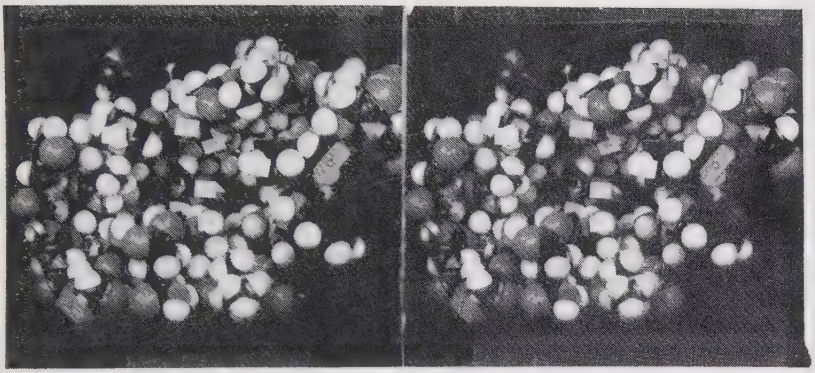
Table 13 presents data on the sulfonamide inhibitors studied in HCAC crystals. Salamide binds with one sulfonamide group to the zinc ion and with the other sulfonamide group to His 63 (3). It is more difficult to modify His 63 with bromopyruvate when salamide is bound

to the enzyme (108), which is to be expected due to the binding to His 63. In the HCAC-AMSulf complex the sulfonamide group is bound to the zinc ion and the benzene ring stretches out toward Gln 91 and Phe 129.

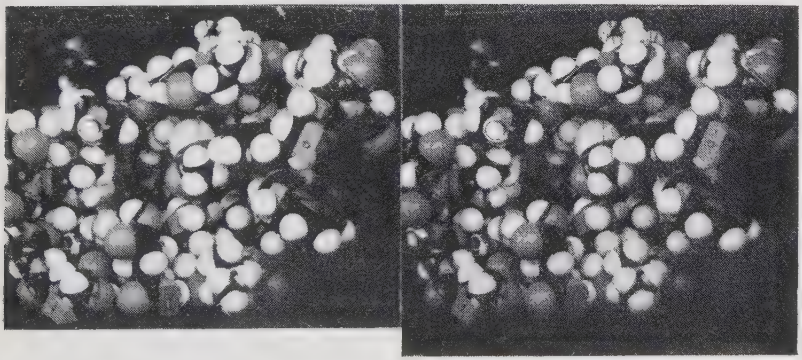
For HCAC-acetazolamide (HCAC-Diamox) 2.5 Å three-dimensional film data were collected. The sulfonamide group is in the same position as for the previous sulfonamide complexes and the distance between the zinc ion and sulfonamide position is given in table 14. The heterocyclic ring stretches out toward the active site opening in almost the same direction as in the AMSulf complex. The acetylamide group is in van der Waals contact with Phe 129, and the oxygen of the group is within hydrogen bond distance of Gln 91 ($N\varepsilon$).

Table 13. Location of various sulfonamides in HCAC.

DERIVATIVE	RESOLUTION (Å)	SYMBOL	BINDING SITE IN THE PROTEIN
Salamide	2.0 (projections only)	SAL_3	Zn, His 63(Gln 91)
AMSulf	2.0 (3-dimensional)	A_1	Cys 204
		A_2	Arg 26, Glu 203
		A_3	Zn, Gln 91, Phe 129
		A_4	His 63
		A_5	His 14
Acetazolamide (DIAMOX)	2.5 (3-dimensional)	X_3	Zn, Gln 91, Phe 129



a



b

Fig. 27 Packing model of a) the active site of HCAC and b) the binding of Diamox in the active site.

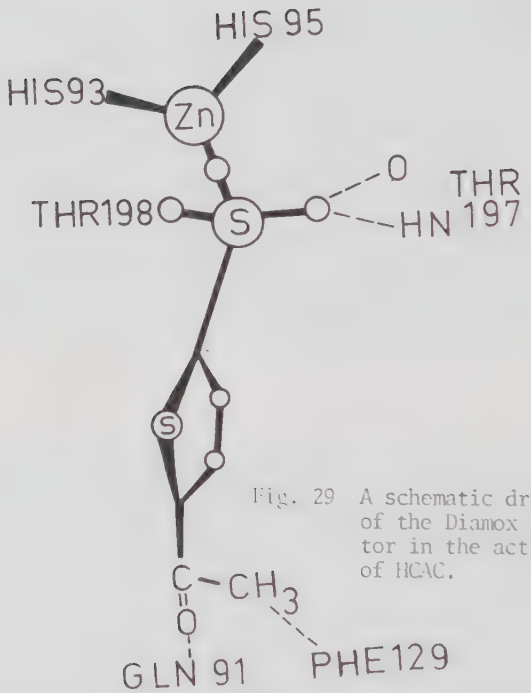


Fig. 29 A schematic drawing of the Diamox inhibitor in the active site of HCAC.

The electron densities of the sulfonamide groups in the HCAC HCAC-AMSulf and HCAC-Diamox complexes indicate that the sulfonamide group is bound to the zinc ion through one of the oxygen atoms or the nitrogen atom. The remaining two atoms are in hydrogen bond distance to Thr 197 (O$_{\gamma}$), the NH group of the same residue and Thr 198 (O$_{\gamma}$). Figures 27 and 28 show the binding of the sulfonamide group and the orientation of the heterocyclic ring in the HCAC-Diamox complex. The planar ring in the AMSulf and Diamox complexes is positioned such that one side is exposed to the hydrophilic and the other side to the hydrophobic side of the active site.

The sulfonamide group is very specifically bound to the protein and one might expect other aromatic sulfonamides to be positioned similarly to the Diamox complex, although the ring may be free to rotate around the S-C bond resulting in different possibilities for hydrogen bonding as in the Salamide complex.

Several methods have indicated that anions compete with sulfonamides in bonding to the enzyme (108, 109, 110, 111) and the results have been consistent with the interpretation that the site which these inhibitors compete for is close to the metal. The structure analyses support this because the sulfonamides and anions bind at the same site.

Table 14. Distances between the zinc ion and the sulfonamide position for certain sulfonamides.

Salamide	AMSulf	Acetazolamide
2.9	3.0	3.0

STRUCTURE STUDIES AT LOW pH

Interesting suggestions have been made on the catalytic mechanism (112, 113). The essential feature in these mechanisms is a zinc bound hydroxyl ion. At the present stage there is not enough data to propose a completely unambigous catalytic mechanism. To solve the mechanism it is necessary to know the substrate binding site. One way to try to find the carbon dioxide binding site is to study crystals of the native enzyme at low pH where the activity is decreased and the solubility of the substrate increased. Crystals were obtained at low pH (6, 0) by soaking pH 8.5 enzyme crystals in

2.3 M ammoniumsulfate and Bis Tris buffer, and lowering the pH stepwise to pH 6.0. The high pH and low pH crystals were isomorphous to high resolution, and the three-dimensional data collection using these crystals is in progress.

SUMMARY

The HCAC molecule is very irregular in shape with overall dimensions 40x42x55 Å. The active site is a 15 Å deep cavity at the bottom of which the essential zinc ion is located. The main chain is well defined and various side chains can easily be identified. About 20 % of the residues are in seven helical segments and 35 % are involved in β-structures. One large pleated sheet runs through the whole molecule and involves 20 % of the residues. The metal ion is bound to the pleated sheet by three histidyl ligands arising from two adjacent chains. The fourth ligand is a water molecule or a hydroxyl ion and the metal coordination is closest to tetrahedral.

In carbonic anhydrase in which the zinc ion has been replaced by other divalent metals, except for Hg(II), the metal ion is located in the same position as in the zinc enzyme. In the $HgCl_2$-substituted HCAC the metal is displaced 0.6 Å relative to the zinc ion. This displacement is in the same direction as in the isozyme HCAB where the Hg(II) is bound to Thr 197 and has an octahedral coordination.

Anions and aromatic sulfonamides are inhibitors of carbonic anhydrases. The anion and the sulfonamide group of the aromatic sulfonamides bind at the same site which is near the metal ion in a pocket formed by residues Thr 197, Leu 196 and Trp 207. There is no solvent molecule between the anion and the metal ion. The structure result indicate that the larger the halide ion is the more contacts does it have with the protein and thus the larger halide ions are better inhibitors than the smaller ones.

The fact that sulfonamides are such powerful inhibitors of carbonic anhydrase seems be due to the specific binding of the sulfonamide group to the protein via Thr 197, Thr 198 and the zinc ion. Their aromatic groups interact with both the hydrophilic and the hydrophobic side of the active site, which would further enhance the binding of these inhibitors. The differences in bidning strengths between various cyclic sulfonamides are probably also due to differences in hydrogen bonding of the aromatic moieties (to His 63, Gln 91 and Glu 68).

Anions bind to two regions in the HCAC molecule; one site is in the active site and the other on the surface. In the presence of MMTGA the site on the surface disappears, a new site appears at His 63, and the anions in the active site are somewhat closer to the zinc ion.

The comparison of the structures of HCAC and the isoenzyme HCAB indicate that, the main chain folding is almost identical, the aromatic clusters are almost identically located, and most of the residues in the active site are the same, the zinc environment with the hydrogen bonding system around His 118 is identical. It seems likely that the same residues are involved in the catalytic mechanism in the two isoenzymes. The following groups are very likely of importance for the reaction mechanism.

1) The zinc ion with the distorted tetrahedral coordination
2) The peculiar binding of His 93
3) Amino acids Thr 197, Gln 91 and His 63
4) The hydrogen bonding system around His 118

ACKNOWLEDGEMENTS

This work has been carried out in the Biological structure group, The Wallenberg Laboratory, Uppsala University, Uppsala.

I am indebted to my group leader Professor B. Strandberg and to Dr K.K. Kannan for their guidance throughout the work.

I would like to thank my co-workers S. Lövgren, M. Petef, A. Ohlsson and K. Fridborg for valuable assistance.

I am greatly thankful to Ing. V. Klimecki for constructing our automatic microdensitometer.

I would also like to thank Dr B. Notstrand, Dr A.Liljas and Dr P.Lentz for fruitful discussions.

I am very greatful to Dr O. Mårtensson for placing Courtould atomic models at our disposal.

I am also very greatful to Professor A.Fredga and to Mr T.Unge for preparation of heavy atom derivatives.

I am indebted to Professors I.Olovsson and C.Kurland for the facilities placed at our disposal.

For stimulating discussions during his visit to our laboratory I would like to thank Professor M.F. Perutz.

I am very greatful to P. Krantz, C.Wallsten and L.G.Pettersson for technical assistance.

I have very much appreciated the helpful cooperation of the whole staff at the Uppsala University Data Center.

I am deeply indebted to G.Lindman and G.Johansson for typing the manuscript, H.Ukkonen for the photographical work and Dr P.Lentz for correcting the language.

For valuable discussions I would like to thank Drs P.O.Nyman, S.Lindskog, L.Hendersson, H.Steiner, L.Strid and D.Hendriksson.

I warmly thank my wife and children for their kind-hearted support of my research work.

I would also like to acknowledge Akademiska Sjukhuset, Uppsala for providing us with human erythrocytes and Pharmacia Fine Chemicals Ltd for generous gifts of Sephadex.

This work was supported by grants from the Faculty of Science, University of Uppsala; the National Institutes of Health, U.S. Public Health Service (Grant No. AI 07382); the Swedish Computing Authorities; the Swedish Medical Research Council (Grant No. B73-13X-26-09); the Swedish Natural Science Research Council (Grant No. 2142); the Tricentennial Fund of the Bank of Sweden and the Knut and Alice Wallenberg Foundation.

REFERENCES

1. Kannan, K.K., Liljas, A., Vaara, I., Bergstén, P.C., Lövgren, S., Strandberg, B., Bengtsson, U., Carlbom, U., Fridborg, K., Järup, L., and Petef, M. Cold Spring Harbor Symposia on Quantitative Biology 36 (1971) 221.

2. Liljas, A., Kannan, K.K., Bergstén, P.C., Vaara, Fridborg, K., Strandberg, B., Carlbom, U., Järup, L., Lövgren, S., and Petef, M. Nature New Biology 235 (1972) 131.

3. Bergstén, P.C., Vaara, I., Lövgren, S., Liljas, A., Kannan, K.K., and Bengtsson, U. Alfred Benzon Symposium IV (Rörth, M. and Astrup, D. eds.) Munksgaard p. 363 (1971).

4. Vaara, I., Lövgren, S., Liljas, A., Kannan, K.K., and Bergstén, P.C. The Second International Conference on Red Cell Metabolism and Function (Brewer, J. ed.) Ann Arbor, Vol. 28 (1972) p. 169.

5. Libscomb, W.N., Hartscuk, J.A., Reeke, G.N., Quiocho, F.A., Bethge, F.H., Ludwig, M.L., Steitz, T.A., Muirhead, H., and Coppola, J.C., Brookhaven Symp. Biol. 21, 24 (1968).

6. Koshland, D.E. Jr. The Enzymes (Boyer, P.D. ed.) Vol. I, p. 342.

7. Kretzinger, R.H. Nature New Biology Vol. 240 Nov. 15 (1972).

8. Poljac, R.J., Amzel, L M , Avey, H.P., Chen, B.L., Phizackerley, R.P., and Saul, F. Proc.Nat.Acad.Sci. Vol. 70, No. 12 p. 3305.

9. Veitch, F.P., and Blankenship, L.C., Nature 197, 76 (1963).

10. Meldrum, N.U., and Roughton, F.J.W., J.Physiol. (London) 75, 4P and 15P (1932).

11. Stadie, W.C., and O'Brien, H. IBC 100, (1933); 103, 521 (1933).

12. Keilin, D., and Mann, T. Nature 144, 442 (1939).

13. Adler, L., Brundell, J., Falkbring, S.O., and Nyman, P.O., Biochim.Biophys.Acta 284, 298 (1972).

14. Tobin, A.J., J.Biol.Chem. 245, 2656 (1970).

15. Rossi, C., Chersi, A., and Cortivo, M., Chemical, Biochemical and Physiological Aspects, NASA SP-188, Washington, D.C., 1969, p.131.

16. Pocker, Y., and Ng, Joan S.Y., Biochemistry, Vol. 12, No 25, (1973).

17. Derrien, Y., and Laurent, G., Exposes Ann.Biochim.Med. 29, 167 (1969).

18. Lindskog, S., and Malmström, B.G., Biochem.Biophys.Res. Commun. 2, 213, (1960).

19. Maren, T.H., Physiol.Rev. 47, 595 (1967).

20. Tashian, R.E., Tanis, R.J., and Ferrel, R.E., in Alfred Benzon Symp. IV, M.Rörth and P. Astrup (Eds.), New York, Academic Press, 353 (1972).

21. Kalifah, R.G., J.Biol.Chem. 246, 2561 (1971).

22. Pocker, Y. and Meany, J.E., Biochemistry 4, 3535 (1965).

23. Strandberg, B., Tilander, B., Fridborg, K., Lindskog, S., and Nyman. P.O., J.Mol.Biol. 583 (1962).

24. Tilander, B., Strandberg, B., and Fridborg, K., J.Mol.Biol. 12 740 (1965).

25. Bror Strandberg, Arkiv för kemi , Band 28, 1 (1967).

26. Fridborg. K., Kannan, K.K., Liljas, A., Lundin, J., Strandberg,B., Strandberg, R., Tilander, B., and Wirén, G., J.Mol.Biol., 25, 505 (1967).

27. Nyman, P.O., Biochim.Biophys.Acta 52, 1 (1961).

28. Armstrong, J.Mc D., Myers, D.V., Verpoorte, J.A., and Edsall, J.T., J.Biol.Chem. 241, 5137 (1966).

29. Schneider, F., and Liefländer, M., Z.Physiol.Chem. 334, 279 (1963).

30. Klimecki. V., private communication.

31. Dickerson, R.E., Kendrew, J.C., and Strandberg, B.E., Acta Cryst. 14, 1188 (1961).

32. Muirhead, H., Cox, J.M., Mazzarella, L., and Perutz, M.F., J. Mol.Biol. 28, 117 (1967).

33. Blow, D.M., and Crick, F.H.C., Acta Cryst. 12, 794 (1959).

34. Richards, F.M., J.Mol.Biol. 37, 225 (1968).

35. Henderson, L.E., Henriksson, D., Nyman, P.O., and Strid, L., in Alfred Benzon Symp. IV, M. Rörth and P. Astrup (eds.), New York, Academic Press, 341 (1972).

36. Henderson, L.E., Henriksson, O., and Nyman, P.O., Biochem.Biophys. Res. Commun. 52, 1388 (1973).

37 a Wyckoff, H.W., Hardman, K.D., Allewell, N.M., Inagami, T., Johnson, L.N., and Richards, F.M. J.Biol.Chem. Vol 242, No 17, p. 3984 (1967).

37 b Wright, C.S., Alden, R.A., and Kraut, J., Nature 221, 235 (1969).

38. Kartha, G., and Ambady, G., Nature Vol 247 Jan. 25 (1974).

39 a Torchia, D.A., Biochemistry, Vol 11 No 8 (1972).

39 b Kendrew, J.C., Brookhaven Symposia in Biology No 15 (1962).

40. Richards, F.M., J.Mol.Biol. 37, 225 (1968).

41. Ramakrishman, C., and Ramachandran, G.N., Biophys. J. 5, 909 (1965).

42. Crawford, J.L., Lipscomb, W.N., and Schellman, C.G., Proc.Nat. Acad.Sci. USA 70, 538 (1973).

43. Venkatachalam, C.M., Biopolymers 6, 1425 (1968).

44. Henderson, L.E., in CO_2: Chemical, Biochemical and Physiological Aspects (R.E. Forster, J.T. Edsall. A.B. Otis, and F.J.W. Roughton, Eds) National Aeronautics and Space Administration, Washington, D.C., 1969, p. 121.

45. Strid, L. Private communication.

46. Carlsson, U., Henderson, L.E., and Lindskog, S. Biochim.Biophys. Acta, 310, 376 (1973).

47. Tanford, C, Aune, K.C., and Ikai, A. J.Mol.Biol. 73, 185 (1973).

48. Yazgan, A., and Henkens, R.W., Biochem 11, 1314 (1972).

49. Notstrand, B. Personal communication.

50. Riddiford, L.M., Stellwagen, R.H., Mehta, S., and Edsall, J.T., J.Biol.Chem. 240, 3305 (1965).

51. Nilsson, A., and Lindskog, S., European J.Biochem. 2, 304 (1967).

52. Riddiford, L.M., J.Biol.Chem. 239, 1079 (1964).

53. Campbell, J. Personal communication.

54. King, R.W., and Roberts, G.C.K., Biochem., 10, 558, (1971).

55. See reference 50.

56. See reference 50.

57. Verpoorte, J.A., and Lindblow, C., J.Biol.Chem. 243, 5993 (1968).

58. Riddiford, L.M., J.Biol.Chem. 240, 168 (1965).

59. Beychok, S., Armstrong, J.McD., Lindblow, C., and Edsall, J.T. J.Biol.Chem. 241, 5150 (1966).

60. Tashian, R. Personal communication.

61. Kandel, M. Personal communication.

62. Timasheff, S.N., Susi, H., and Stevens, L., J.Biol.Chem. 242, 5467 (1967).

63. Coleman, J.E., Biochemistry 4, 2644 (1965).

64. Myers, D.V., and Edsall, J.T., Proc.Nat.Acad.Sci. 53, 169 (1965).

65. Blundell, T.L., Dodson, G., Hodgin, D, and Mercola, D., Adv.Prot. Chem. 26, 279 (1972).

66. Coleman, P.M., Jansonius, J.N., and Matthews, B.W., J.Mol.Biol. 70, 701 (1972).

67. Freeman, H.C. in C.B. Anfinsen, Jr., M.L. Anson, J.T. Edsall, and F.M. Richards (Eds) Adv.Prot.Chem. 22, 257 (1967).

68. Lindskog, S., J.Biol.Chem. 238, 945 (1963).

69. Williams, R.J.P., Biopolymers Symposia 1, 515 (1964).

70. Lindskog, S., and Nyman, P.O., Biochim.Biophys.Acta 85, 462 (1964).

71. Dennard, A.E., and Williams, R.J.P., in Transition Metal Chemistry (R.L. Carlin. Ed.) 2, 115 (1966).

72. Lindskog, S., and Ehrenberg, A., J.Mol.Biol. 24, 133 (1967).

73. Coleman, J.E., J.Biol.Chem. 242, 5212 (1967).

74. See reference 73.

75. Cockle, S.A. Private communication.

76. Taylor, J.S., Mushak, P., and Coleman, J.E., Proc.Nat.Acad.Sci. U.S. 67, 1410 (1970).

77. Tilander, B., Strandberg, B, and Fridborg, K., J.Mol.Biol. 12, 740 (1965).

78. Carlsson, U. Private communication.

79. Kandel, S.I., Wong, S-C.C., Kandel, M., and Gornall, A. J. Biol.Chem. Vol.243 No 9, 2237 (1968).

80. Göthe, P.O., and Nyman, P.O., Febs Letters, Vol.21, No 2 (1972).

81. Göthe, P.O. Inaugural dissertation, Göteborg 1972.

82. Fabry, M.E., Koening, S.H., and Schillinger, W.E., J.Biol.Chem. 245, 4256 (1970).

83. Verpoorte, J.A., Metha, S., and Edsall, J.T., J.Biol.Chem. 242, 4221 (1967).

84. See reference 73.

85. Donohue, J., Proc.Nat.Acad.Sci.U.S. 39, 470 (1953).

86. Pauling, L., and Corey, R.B., Proc.Nat.Acad.Sci. U.S. 37, 729 (1951).

87. Wyckoff, H.W., Tsernoglou, D., Hanson, A.W., Knox, J.R., Lee, B., Richards, F.M. J.Biol.Chem. 245, 305 (1970).

88. Drenth, J., Jansonius, J.N., Koekoek, R., Sluyterman, L.A.A., and Wolters, B.G., Phil.Trans.Roy.Soc. Lond. B. 257, 23 (1970).

89. Adams, M., Ford, G.C., Koekoek, R., Lentz, P.J., McPherson, A., Rossmann, M.G., Smiley, I.E., Schevitz, R.W., and Wonacott, A.J., Nature 227, 1098 (1970).

90. Reeke, G.N., Hartsuck, J.A., Ludwig, M.L., Quiocho, F.A., Steitz, T.A., and Lipscomb, W.N., Proc.Nat.Acad.Sci.U.S. 58, 2220 (1967).

91. Lindskog, S., Biochemistry 5, 2641 (1966).

92. Hofmeister, F., Arch.Experimental Pathol.Pharmacol. 24, 247 (1888).

93. Verpoorte, J.A., Metha, S., and Edsall, J.T., J.Biol.Chem. 242, 4221 (1967).

94. Kernohan, J.C., Biochim.Biophys.Acta 96, 304 (1965).

95. Pocker, Y., and Stone, J.T., J.Am.Chem.Soc. 87, 5497 (1965).

96. Pocker, Y., and Stone, J.T., Biochemistry 6, 668 (1967).

97. Ward, R.L., Biochemistry 8, 1879 (1969).

98. Zeppezauer, M., Lindman, B., Forsén, S., and Lindqvist, I., Biochem.Biophys.Res.Commun. 37, 137 (1969).

99. Ward, R.L., Biochemistry, 9, 2447 (1970).

100. See reference 80.

101. Bar, D., Actualites Pharmacol. 15, 1 (1963).

102. Maren, T.H., Physiol. Rev. 47, 595 (1967).

103. King, R.W., and Burgen, A.S.V., Biochim.Biophys. Acta 207, 278 (1970).

104. Taylor, P.W., King, R.W., and Burgen A.S.V., Biochem. 9, 2638 (1970).

105. Taylor, P.W., King, R.W., and Burgen, A.S.V., Biochem. 9, 3894, (1970).

106. See reference 101.

107. See reference 102.

108. Liljas, A., "Crystal Structure Studies of Human Erythrocyte Carbonic Anhydrase C at High Resolution", Almqvist & Wiksell, Stockholm (1971).

109. Coleman, J.E., Nature 214, 193 (1967).

110. Ward, R.L., and Fritz, K.J., Biochem.Biophys.Res.Commun. 39, 707 (1970).

111. Lindskog, S., and Thorslund, A., European J.Biochem. 3, 453 (1968).

112. Khalifah, R.G.,J.Biol.Chem., 246, 2561 (1971)

113. Coleman, J. E. In Inorganic Biochemistry (Eichhorn, Ed.) (1971) In press.